

# **Die Regulation der prokaryontischen Transkription auf molekularer Ebene**

Dissertation

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

Vorgelegt von  
Diplom Biochemiker  
Stefan Josef Prasch

Bayreuth 2008

Die vorliegende Arbeit wurde von September 2004 bis Dezember 2007 am Lehrstuhl für Struktur und Chemie der Biopolymere 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.

Promotionsgesuch eingereicht am: 04.01.2008

Tag des wissenschaftlichen Kolloquiums: 10.06.2008

Erster Gutachter: Prof. Dr. Paul Rösch  
Zweiter Gutachter: Prof. Dr. Matthias Ullmann  
Vorsitzender: Prof. Dr. Carlo Unverzagt  
Prof. Dr. Stephan Clemens

# Inhaltsverzeichnis

Inhaltsverzeichnis.....	i
Kurzfassung .....	ii
Summary .....	iv
1. Einleitung .....	1
1.1. Der Phage $\lambda$ .....	1
1.1.1. Lytischer Zyklus.....	1
1.1.2. Lysogener Zyklus.....	2
1.2. Überblick über die bakterielle Transkription .....	3
1.2.1. Die RNA-Polymerase.....	3
1.2.2. Die Initiation .....	6
1.2.3. Die Elongation .....	9
1.2.4. Die Termination .....	10
1.3. Antitermination.....	12
1.3.1. Der Antiterminationsmechanismus des Phagen $\lambda$ .....	13
1.3.2. Die Antitermination an <i>rrn</i> Operons .....	16
1.4. Der Elongationsfaktor NusA.....	17
2. Ziele.....	21
3. Zusammenfassung und Diskussion der Ergebnisse.....	22
3.1. Charakterisierung von Proteinkomplexen.....	22
3.2. $\lambda$ N in Komplex mit der carboxyterminalen Domäne AR1 aus NusA.....	23
3.3. NusA-AR2 bindet an die $\alpha$ CTD.....	25
3.4. Vergleich von AR1 und AR2.....	26
3.5. Wechselwirkung zwischen NusA und RNA .....	27
3.6. Aufbau des Antiterminationskomplexes .....	28
4. Abkürzungsverzeichnis .....	30
5. Literaturverzeichnis.....	31
6. Publikationsliste .....	41
6.1. Einzelarbeit A .....	41
6.2. Einzelarbeit B.....	41
6.3. Einzelarbeit C.....	42
6.4. Einzelarbeit D .....	42
Einzelarbeit A .....	43
Einzelarbeit B.....	57
Einzelarbeit C.....	71
Einzelarbeit D .....	87
Danksagung.....	99
Erklärung.....	101

## Kurzfassung

In *Escherichia coli* wird die RNA-Synthese durch eine aus fünf Untereinheiten bestehende RNA-Polymerase (RNAP) katalysiert. Für die Bildung der Phosphodiesterbindung in der RNA sind die  $\beta$ - und  $\beta'$ -Untereinheiten verantwortlich. Die aminoterminalen Domänen der beiden  $\alpha$ -Untereinheiten sind für den korrekten Aufbau zuständig, während die durch einen flexiblen Linker verbundenen carboxyterminalen Domänen ( $\alpha$ CTD) regulatorische Aufgaben erfüllen. Die zuletzt entdeckte  $\omega$ -Untereinheit fördert ebenso den Aufbau, ist aber für das Überleben von *Escherichia coli* nicht notwendig.

Der Ablauf der Transkription wird begrifflich in die Initiation, Elongation und Termination eingeteilt. Bei der Initiation bindet ein so genannter  $\sigma$ -Faktor an die RNAP. Das so entstandene Holoenzym erkennt den Promotor und startet die RNA-Synthese. Der  $\sigma$ -Faktor verlässt die RNAP, es bildet sich der Transkriptionselongationskomplex, der zu einer stark gesteigerten RNA-Syntheserate führt. Im letzten Schritt erfolgt der Kettenabbruch, woraufhin freigewordene RNAP erneut einen Transkriptionszyklus starten kann.

Eine zentrale Rolle bei der Regulation der Elongation und Termination spielt das essentielle NusA Protein (*N utilization substance A*), das im Verbund mit NusB, NusE und NusG, vor allem die Geschwindigkeit der RNA-Synthese steuert. Der Aufbau von NusA aus sechs Domänen spiegelt die vielfältigen Aufgaben des Proteins. Die aminoterminal Domäne vermittelt die Interaktion mit der RNAP. Die RNA-Bindungsregion wird durch die S1-, KH1- und KH2-Domänen aufgebaut. Die beiden carboxyterminalen *acidic repeats* AR1 und AR2 erfüllen jeweils unterschiedliche Aufgaben, die bisher nur zum Teil aufgeklärt sind und im Rahmen dieser Arbeit untersucht werden.

Eine weitere Rolle spielt NusA bei der Antitermination. Dieser zuerst im Phagen  $\lambda$  entdeckte Mechanismus bewirkt, dass die Transkription nicht an einem Terminationssignal abbricht. Das dafür notwendige Antiterminationssignal besteht aus zwei bestimmten RNA-Sequenzabschnitten, *boxA* und *boxB*, die durch eine Trennsequenz miteinander verbunden sind. Während die *boxA* sowohl im Phagen  $\lambda$  als auch in den für die ribosomale RNA codierenden Operons stark konserviert ist, weist die *boxB* in den beiden genannten Fällen nur geringe Gemeinsamkeiten auf. Das NusA Protein bindet an das Antiterminationssignal und an die RNAP. Die so veränderte RNAP ist nun in der Lage, stromabwärts gelegene Terminationssignale zu überlesen.

In dieser Arbeit wurde zuerst die Rolle der AR1 Domäne untersucht. Eine bereits bekannte Kristallstruktur zeigt AR1 in Komplex mit  $\lambda$  N. Jedoch fehlt in dieser Struktur die zweite Hälfte des  $\lambda$  N Bindungsmotives (Aminosäurereste 42 bis 47), ebenso wie AR2, was zur

Vermutung führte, dass diese Domäne  $\lambda$  N(42-47) binden könne. Zur Klärung dieser Frage wurden die beiden Domänen AR1 und AR2 jeweils einzeln kloniert und im Hinblick auf ihre Bindung an das  $\lambda$  N Protein getestet. In dieser Arbeit wurde eine ausschließliche Bindung von AR1 an  $\lambda$  N gefunden, inklusive der in der Kristallstruktur fehlenden Aminosäurereste. Zusätzlich induziert AR1 bei der Bindung eine helikale Konformation im Bereich des Argininreichen Motives von  $\lambda$  N (Aminosäurereste 1 bis 22), wodurch ein theoretisches Modell der  $\lambda$  N Funktionsweise bestätigt wurde.

Zudem wurde die Interaktion von NusA-AR2 charakterisiert. Biochemische Daten ließen auf eine Wechselwirkung von NusA mit  $\alpha$ CTD schließen. Basierend darauf wurde in dieser Arbeit nachgewiesen, dass in diesem Fall ausschließlich AR2 für die Wechselwirkung mit  $\alpha$ CTD verantwortlich ist. Mit Hilfe von NMR-Experimenten wurde im Anschluss daran die Komplexstruktur berechnet. Die resultierende Bindungsfläche seitens  $\alpha$ CTD lieferte dabei Rückschlüsse, wie NusA möglicherweise die RNAP von der Initiations- in die Elongationsphase überführt. Ebenso erlaubten die neu gewonnenen Ergebnisse einen Vergleich der Bindungseigenschaften von AR1 und AR2.

Weiterhin wurde in dieser Arbeit die Wechselwirkung von NusA mit dem Antiterminationssignal studiert. Dazu wurde eine NusA Deletionsmutante kloniert und gereinigt, die ausschließlich aus den bisher vermuteten RNA-Bindungsdomänen S1, KH1 und KH2 bestand. Anhand der gemessenen Fluoreszenzanisotropie-Daten wurde die Bindung im Bereich der Trennsequenz lokalisiert. Die schwächere Affinität des Phagen  $\lambda$  Signals *nutL* und *nutR* im Vergleich zur ribosomalen RNA erklärt die Notwendigkeit des zusätzlichen Faktors  $\lambda$  N im Phagen System.

## Summary

RNA synthesis in *Escherichia coli* is catalyzed by RNA polymerase (RNAP), whose core consists of five subdomains,  $\alpha_2\beta\beta'\omega$ . Whereas the  $\beta$  and  $\beta'$  subunits catalyze RNA polymerization, the two  $\alpha$  subunit N-terminal domains orchestrate RNAP assembly and the C-terminus of the  $\alpha$  subunit ( $\alpha$ CTD) forms an independent domain attached via an unstructured linker. The  $\omega$  subunit is dispensable for cell viability, but may act as a chaperone and promote RNAP assembly *in vitro*.

The transcription process is subdivided into initiation, elongation, and termination. During initiation a so called  $\sigma$ -factor binds the RNAP. The holoenzym recognizes the promoter and thus enables the start of the transcription process. After  $\sigma$ -factor release, the high processive transcription elongation complex is formed. At the end of transcription, chain elongation is interrupted and the released RNAP can restart a new transcription cycle.

The essential NusA protein (N utilization substance A) together with the additional host factors NusB, NusE, and NusG modulates transcription pausing, termination, and antitermination. The multiple functionality of NusA is reflected by its multidomain structural organization. The N-terminal RNAP-binding domain is connected through a flexible hinge helix to the RNA binding interface that is composed of three domains, S1, KH1, and KH2. Finally, NusA has an additional C-terminal extension consisting of two so called *acidic repeats* AR1 and AR2.

NusA plays also an important role in antitermination in the phage  $\lambda$ /*Escherichia coli* system. This mechanism inhibits the termination of transcription at intrinsic and Rho-dependent termination signals. The necessary RNA based signal is composed of *boxA* and *boxB*, separated by a spacer. NusA together with auxiliary host factors binds the antitermination signal and alters RNAP to prevent transcription termination. For phage  $\lambda$  antitermination, an additional  $\lambda$  phage protein called  $\lambda$  N is necessary. The aim of this work has been to elucidate different interactions of NusA within the antitermination complex.

As a first step, the interaction of NusA AR1 and  $\lambda$  N has been investigated. Heteronuclear NMR experiments with  $\lambda$  N in complex with AR1 revealed that upon complex formation the  $\lambda$  N-binding interface, residues 37 to 47, adopt a rigid structure similar to the known X-ray structure. As the second parts of the  $\lambda$  N-binding interface, residues 42 to 47 as well as AR2 are missing in the X-ray structure, this work intended to clarify their role in the antitermination complex. From titration experiments it could be shown that AR2 is not part of the  $\lambda$  N-NusA complex. Furthermore, the missing residues of  $\lambda$  N also interact with AR1. Moreover,

upon binding of NusA induction of a weak helical structure in the  $\lambda$  N *boxB* RNA-binding region, residues 1 to 22, could be demonstrated, in agreement with one of the theoretical models of  $\lambda$  N action.

Furthermore, interactions of the second acidic repeat, AR2, with NusA were studied. From biochemical data an interaction of the C-terminal part of NusA with  $\alpha$ CTD had been proposed. Here we could demonstrate that only AR2 is involved in this complex. Subsequently, the complex structure has been determined using NMR-techniques. AR2 blocks the “265 determinant” on  $\alpha$ CTD which is known to mediate the interaction with an important initiation regulatory element, the *upstream activating element*. Based on these results, the interplay between AR2 and  $\alpha$ CTD could displace this element, suggesting the NusA- $\alpha$ CTD interaction to be a key switch from initiation to elongation.

Finally, this work describes in quantitative terms the interaction of NusA with antitermination signals composed of a *boxA* and *boxB* separated by a spacer. Isotropic and anisotropic fluorescence equilibrium titration revealed that the RNA binding site for NusA is located in the spacer. The lower affinity for the phage  $\lambda$  antitermination signals *nutL* and *nutR* compared to the ribosomal signal in the *rrn* Operon explains the necessity of the auxiliary factor  $\lambda$  N.





# 1. Einleitung

## 1.1. Der Phage $\lambda$

Unser Verständnis der Transkription wurde in den letzten Jahren dramatisch erweitert. Neben zahlreichen neuen genetischen und biochemischen Erkenntnissen gelang es, das wichtigste Enzym der Transkription, die RNA-Polymerase (RNAP), zu kristallisieren und die Struktur zu bestimmen. Den vorläufigen Höhepunkt stellte die Kristallstruktur einer eukaryotischen RNAP (RNAP II) aus Hefe dar (Cramer *et al.*, 2001; Gnatt *et al.*, 2001). Hierfür wurde im Jahr 2006 der Nobelpreis an Roger D. Kornberg verliehen. Der nächste Schritt besteht nun darin, die Regulation der Transkription auf molekularer Ebene zu erforschen. Obwohl die molekularen Mechanismen zur Regulation der Genexpression oft auf relativ einfache Interaktionen zurückgreifen, ist es bereits sehr schwierig, die Gesamtheit regulatorischer Netzwerke in einem kleinen Organismus zu verstehen, so dass das Verständnis von größeren, komplexeren Netzwerken, wie denen des Menschen, derzeit fast unmöglich erscheint. Der „einfache“ Bakteriophage  $\lambda$  ist in der Lage, zwischen zwei alternativen Entwicklungszyklen zu unterscheiden und dient daher oft als Modellsystem. Der Phage  $\lambda$  kann sich auf der einen Seite in das Bakterium *Escherichia coli* (*E. coli*) einschleusen, um sich darin zu vermehren (lytischer Zyklus), wobei durch die Lyse der Zelle die gebildeten Phagenpartikel freigesetzt werden. Auf der anderen Seite kann sich der Phage  $\lambda$  in das Wirtsgenom der infizierten Zelle integrieren. Die Fähigkeit zur Vermehrung und Lyse bleibt weiterhin bestehen (lysogener Zyklus) (Court *et al.*, 2007; Friedman und Court, 2001; Gottesman und Weisberg, 2004).

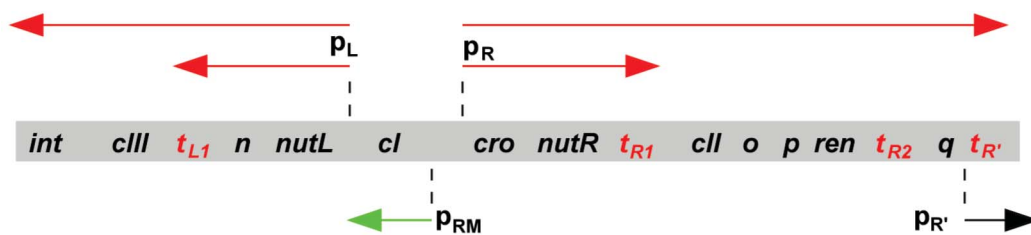
### 1.1.1. Lytischer Zyklus

Im lytischen Zyklus werden hunderte von Phagen freigesetzt. Zuerst wird vom rechten und linken Promotor ( $p_L$  und  $p_R$ ) das Phagen genom bis zu den entsprechenden Terminationsstellen  $t_R$  und  $t_L$  transkribiert (Abb. 1-1; Court *et al.*, 2007). Das als erstes entstandene  $\lambda$  N Protein moduliert die RNAP, die nun in der Lage ist, die Terminationsstellen  $t_R$  und  $t_L$  zu überlesen. Dieser Prozess wird als Antitermination bezeichnet (Abschnitt 1.3.1). Gene stromabwärts von  $t_R$  und  $t_L$ , kodieren verschiedene Funktionen für den lytischen und lysogenen Zyklus. So beinhaltet das  $p_L$  Operon wichtige Gene, um durch Rekombination die Phagen DNA in das Wirtsgenom zu integrieren. Das  $p_R$  Operon enthält neben den Genen *o*, *p* und *ren*, die für die Repli-

kation verantwortlich sind, auch das regulatorische Gen  $q$ . Ähnlich wie  $\lambda$  N bewirkt  $\lambda$  Q das Überlesen der  $t_R$  Terminationsstelle, wodurch die restlichen Gene für den lytischen Zyklus exprimiert werden. Die nun stromabwärts von  $p_R$  transkribierten Gene, codieren Proteine, um die Zellen zu lysieren (Court *et al.*, 2007).

### 1.1.2. Lysogener Zyklus

Abhängig von den Wachstumsbedingungen der infizierten Zelle wird der lysogene Zyklus aktiviert. In diesem Fall wird das Phagenom in das Wirtsgenom integriert und die meisten der viralen Phagenproteine werden abgeschaltet, so dass  $\lambda$  als Prophage fortbesteht. Die ersten Schritte unterscheiden sich dabei nicht vom lytischen Zyklus (Dodd *et al.*, 2005). Nach erfolgter  $\lambda$  N Expression werden die Gene  $int$ ,  $cII$  und  $cIII$  exprimiert (Abb. 1-1). Das Protein Int erlaubt die Integration des Phagenoms in das bakterielle Chromosom. Die regulatorischen Gene  $cII$  und  $cIII$  werden benötigt, um das Gen  $ci$  zu exprimieren. Das  $cII$  Gen befindet sich zwischen dem Terminator  $t_{R1}$  und den Genen für die Replikation und wird somit mit den frühen lytischen Genen exprimiert. Das  $cIII$  Gen liegt zwischen  $t_{L1}$  und dem  $p_L$  Operon. Beide Genprodukte werden ausschließlich dazu benötigt, das  $ci$  Gen zu aktivieren. Das gebildete CI Protein unterdrückt die beiden frühen Promotoren  $p_L$  und  $p_R$ , wodurch die Expression der lytischen Gene verhindert wird (Ptashne, 2006). Sobald das CI Protein gebildet ist, werden die Gene  $cII$  und  $cIII$  nicht weiter exprimiert, da das CI Protein einer Selbstregulation unterliegt.



**Abb. 1-1: Ausschnitt aus dem Genom des Phagen  $\lambda$**

Die frühen Transkripte für  $p_L$  und  $p_R$  Promotoren sind als rote Pfeile gekennzeichnet. Das späte Transkript ab  $p_{R'}$  ist als schwarzer Pfeil dargestellt. Gene sind schwarz und die Terminatoren sind rot gekennzeichnet. Zuerst werden die Transkripte stromabwärts von  $p_L$  und  $p_R$  abgelesen. Für den lytischen Zyklus müssen die Gene stromabwärts von  $p_{R'}$  transkribiert werden, wohingegen für den lysogenen Zyklus die Transkription des CI Proteins stromabwärts von  $p_{RM}$  wichtig ist.

## 1.2. Überblick über die bakterielle Transkription

Der wichtigste Schritt zur Regulation der Transkription des Bakteriophagen  $\lambda$  ist die Antitermination. Um diese verstehen zu können, soll zunächst die Transkription näher erläutert werden. Sowohl die prokaryotische als auch die eukaryotische Transkription können in drei unterschiedliche Schritte eingeteilt werden: Initiation, Elongation und Termination. Da die RNAP das wichtigste Enzym der Transkription darstellt, wird dieses zunächst genauer beschrieben.

### 1.2.1. Die RNA-Polymerase

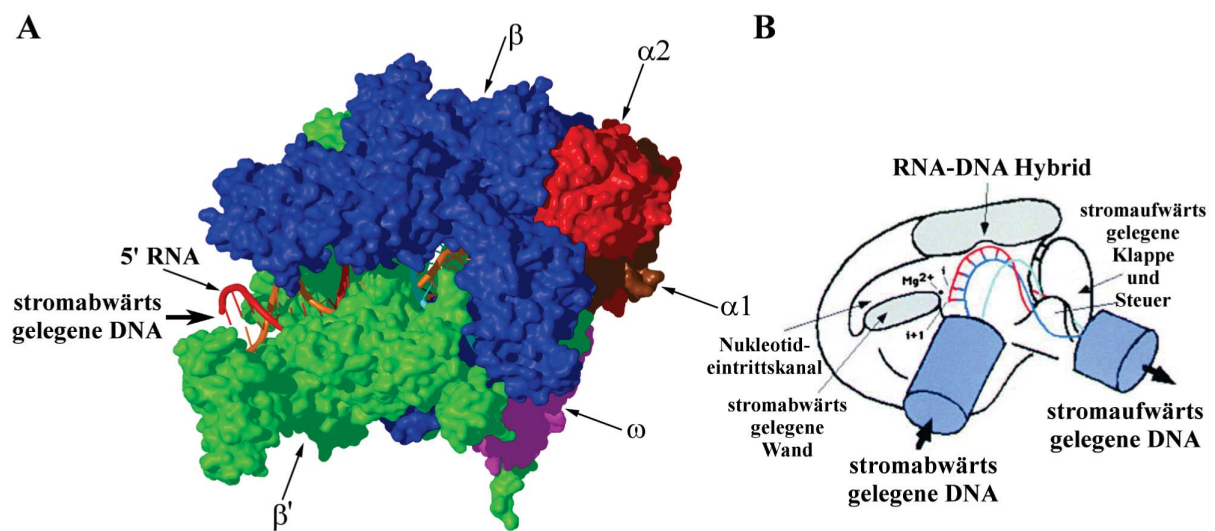
In allen Organismen basiert die Transkription auf einer DNA-abhängigen RNAP (Murakami und Darst, 2003; Shilatifard *et al.*, 2003). Im Gegensatz zu Eukaryonten gibt es bei Prokaryonten nur eine RNAP. Das Herzstück der prokaryotischen *E. coli* RNAP besteht aus fünf Untereinheiten mit der Zusammensetzung  $\alpha_2\beta\beta'\omega$  mit einer Molekülmasse von ca. 400 kDa (Severinov, 2000). Die  $\beta$ - und  $\beta'$ -Untereinheiten katalysieren die RNA-Polymerisation (Abb. 1-2). Die beiden aminoterminalen Domänen der  $\alpha$ -Untereinheit sind für den korrekten Aufbau der RNAP notwendig. Die in allen bisher bestimmten Kristallstrukturen nicht sichtbare carboxyterminale Domäne der  $\alpha$ -Untereinheit ( $\alpha$ CTD) ist über einen flexiblen Linker an die RNAP gebunden (Jeon *et al.*, 1997). Vor allem regulatorische Proteine interagieren mit  $\alpha$ CTD (Abschnitt 1.4). Die  $\omega$ -Untereinheit ist für die Funktionsfähigkeit der RNAP nicht notwendig, unterstützt aber *in vitro* als Chaperon den Zusammenbau der RNAP (Mukherjee und Chatterji, 1997; Mukherjee *et al.*, 1999).

Für den Start der Transkription muss zusätzlich noch ein Initiationsfaktor gebunden werden, der als  $\sigma$ -Faktor bezeichnet wird. Während der Initiation erkennt die RNAP den Promoter, schmilzt die DNA in der Nähe des Startpunktes (bezeichnet als  $i+1$ ) auf und startet die RNA-Synthese. Die verwendeten Nukleotidtriphosphate werden von der RNAP zum einen als Primer zum anderen als Substrat verwendet. Während in den ersten Schritten der Transkription nur kurze Abbruchfragmente entstehen, werden in der Elongationsphase RNA-Ketten polymerisiert, die mehr als 1000 Nukleotide umfassen. Die Polymerisationsgeschwindigkeit beträgt 60-80 Nukleotide pro Sekunde (Gotta *et al.*, 1991; Vogel und Jensen, 1994).

In letzter Zeit wurden Kristallstrukturen von RNAPs aus verschiedenen Organismen gelöst, sowohl als *core* RNAP als auch als Holoenzym (Cramer *et al.*, 2001; Gnatt *et al.*, 2001; Vas-

sylyev *et al.*, 2007a; Zhang *et al.*, 1999). Die Form und Gestalt der „core“ *Thermus thermophilus* (*T. thermophilus*) RNAP korrespondiert sehr gut mit der *E. coli* RNAP, von der bisher nur die Oberflächenkonturen mit Röntgenkristallographie bestimmt werden konnten (Darst, 2001). Somit kann die *T. thermophilus* Struktur als Homologiemodell für Studien der *E. coli* RNAP verwendet werden.

Die kürzlich gelöste Kristallstruktur eines Transkriptionselongationskomplexes (TEC) zeigt die RNAP aus *T. thermophilus* in der Elongationsphase (Vassylyev *et al.*, 2007a; Vassylyev *et al.*, 2007b). Die hochaufgelöste Struktur (2.5 Å) zeigt die exakte Länge des DNA-RNA Hybrides, den Mechanismus der DNA Translokation sowie die Konfiguration der stromabwärts (*downstream*) gelegenen DNA (Abb. 1-2).



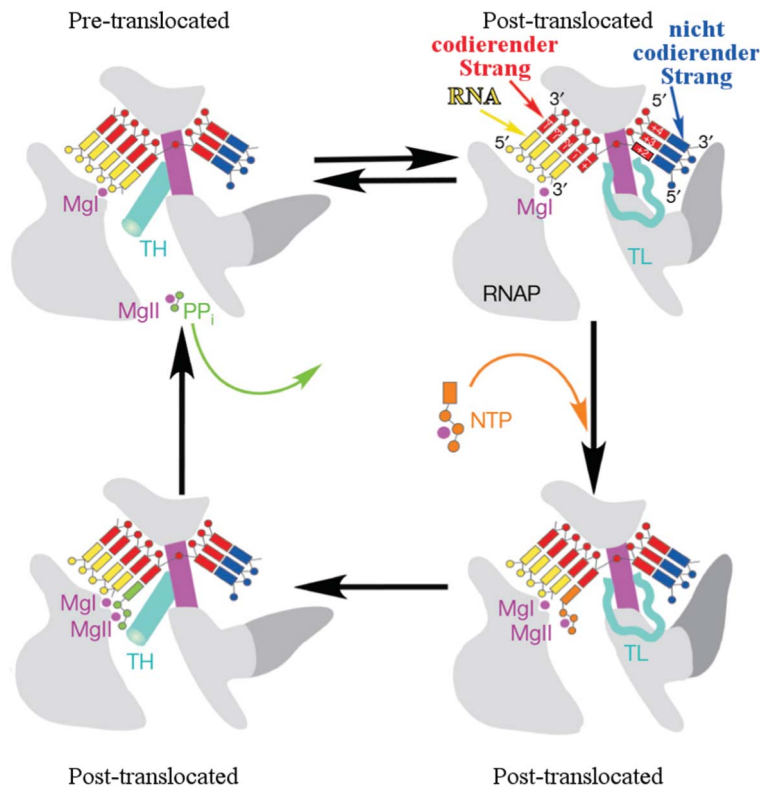
**Abb. 1-2: Struktur der *T. thermophilus* RNAP**

(A) Die Oberflächendarstellung zeigt die beiden  $\alpha$ -Untereinheiten (rot und braun), die  $\beta$  Untereinheit in blau, die  $\beta'$  Untereinheit in grün, sowie die  $\omega$  Untereinheit in magenta. Zusätzlich sind Teile des RNA-DNA Hybrides zu sehen (RNA rot, bzw. DNA orange) (PDB-Kennung 2O5I; Vassylyev *et al.*, 2007a). (B) Die schematische Darstellung der RNAP zeigt die um fast 90° geknickte RNA (Zhang *et al.*, 1999).

Die RNAP besitzt die Form einer Kresschere, deren beide Zangen durch einen tiefen Kanal getrennt sind. Die obere Zange wird hauptsächlich von der  $\beta$ -Untereinheit (blaue Untereinheit in Abb. 1-2 A), die untere von der  $\beta'$ -Untereinheit gebildet (grüne Untereinheit in Abb. 1-2 A). Der Kanal hat einen Durchmesser von  $\sim 27$  Å, der groß genug ist, um die doppelsträngige DNA aufzunehmen (Abb. 1-2 A). Im aktiven Zentrum befinden sich zwei funktionell wichtige  $Mg^{2+}$ -Ionen, die  $\sim 3,9$  Å voneinander entfernt sind und vom Phosphatrückgrat des Substrates koordiniert werden. Das erste  $Mg^{2+}$  wird darüber hinaus durch drei Aspartatreste gebunden, während das zweite  $Mg^{2+}$  von nur einem Aspartatrest komplexiert wird. Zusätzlich wird das zweite  $Mg^{2+}$  über zwei Wassermoleküle komplexiert, zwei weitere Koordinationsplätze

können aber in der Struktur nicht gefunden werden. Das  $Mg^{2+}$  zusammen mit den beiden  $\beta$ - und  $\beta'$ -Untereinheiten katalysiert die Bildung der Phosphodiesterbindung zwischen dem freien 3'-OH der bereits vorhandenen RNA-Kette und dem nächsten mit der DNA-Matrize basengepaarten Ribonukleotid. Der Nukleotid-Eintrittskanal liegt in der Nähe der *dw*DNA und wird durch eine stromabwärts gelegene Wand von dieser getrennt (Abb. 1-2 B). Die DNA-Matrize zeigt einen scharfen Knick ( $\sim 90^\circ$ ) zwischen der stromaufwärts und stromabwärts gelegenen DNA (Abb. 1-2 B). Die einzelsträngige RNA verlässt die RNAP durch den RNA-Austrittskanal, das Phosphodiesterückgrat der RNA nimmt darin die A-Form an. Die *dw*DNA liegt in der B-Form vor, die am Knick im Zentrum der RNAP deformiert ist (Abb. 1-2 B; Vassylyev *et al.*, 2007a). Um eine hohe Mobilität und Prozessivität zu gewährleisten, zeigt die *dw*DNA nur sehr wenige Kontakte mit dem aus  $\beta$ - und  $\beta'$ -Untereinheiten gebildeten Kanal. Im Gegensatz dazu wird das RNA-DNA Hybrid über polare und van-der-Waals-Wechselwirkungen fest vom aktiven Zentrum gebunden. Das stromaufwärts gelegene Ende des RNA-DNA Hybrides wird durch eine als Klappe bezeichnete Schleife getrennt. Zusätzlich ist diese Schleife an der Ausbildung des RNA-Austrittskanals beteiligt. Eine als Steuer bezeichnete weitere Schleife stabilisiert das RNA-DNA Hybrid (Abb. 1-2 B). Die sogenannte Klappendomäne der  $\beta'$ -Untereinheit besitzt eine flexible (*flap-tip*) Helix. Diese Klappenhelix bedeckt den RNA-Austrittskanal und ist ein möglicher Wechselwirkungspartner für verschiedene regulatorische Signale, welche die Termination steuern (Abschnitt 1.2.4; Touloukhonov und Landick, 2003).

Die Struktur zeigt innerhalb der DNA-Matrize exakt eine ungepaarte Base, die mit einem eintretenden Substrat eine Watson-Crick Basenpaarung ausbilden kann (Vassylyev *et al.*, 2007b). Anhand der Struktur und verschiedener biochemischer Versuche wird ein Zweizustandsmodell postuliert (Vassylyev *et al.*, 2007b). Vor dem Beladen der Nukleotidbindungsstelle herrscht ein thermisches Gleichgewicht zwischen dem prä- und dem post-translozierten Zustand (Abb. 1-3). Zuerst bindet ein Nukleotid an die offene Bindungsstelle und stabilisiert so den post-translozierten Zustand. Im weiteren Verlauf wird eine bestimmte Schleife (*trigger loop*) in eine Helix (*trigger helix*) umgefaltet. Dadurch wird das neu einzubauende Nukleotid exakt positioniert. Zudem schließt sich das aktive Zentrum, um nun die Bildung der Phosphodiesterbindung an der RNA zu katalysieren. Die *trigger helix* wird dadurch destabilisiert und faltet sich nun wieder in den *trigger loop* zurück (Abb. 1-3). Die Entfaltung der Helix bewirkt ein Weiterschieben der DNA-Matrize, um einen neuen Polymerisationszyklus starten zu können.



**Abb. 1-3: Nukleotideinbau**

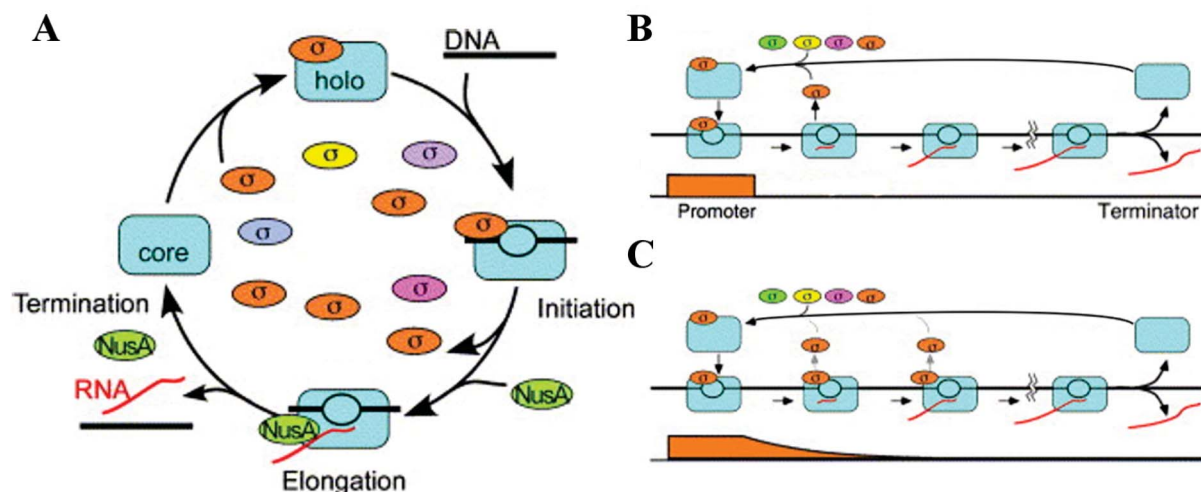
Der Substratbeladungsweg in einer bakteriellen RNAP (grau). Aus der Kristallstruktur kann ein Zweizustandsmodell abgeleitet werden (Vassylyev *et al.*, 2007b). Ein charakteristisches Merkmal des Einbauzyklus ist die Umfaltung des „trigger Loop“ (TL) in die „trigger Helix“ (TH), wodurch eine exakte Positionierung des einzubauenden Nucleotids erfolgt (aus Vassylyev 2007).

### 1.2.2. Die Initiation

Von allen drei Schritten der Transkription ist die Initiation am genauesten reguliert. Durch das Binden des  $\sigma$ -Faktors an die *core*-RNAP wird das Holoenzym gebildet, das nun entlang der DNA die entsprechende Promotor-Erkennungssequenz sucht. Im nächsten Schritt wird die DNA aufgeschmolzen und bildet die Transkriptionsblase. Kurze RNA-Abbruchfragmente werden nun so lange synthetisiert, bis ein erstes stabiles 8 bis 9 Basenpaare (bp) langes RNA-DNA Hybrid gebildet wurde. Danach verlässt der  $\sigma$ -Faktor die RNAP. Die Initiation ist damit abgeschlossen. Der dissoziierte Initiationsfaktor kann nun wieder eine freie RNAP binden und so einen neuen Transkriptionszyklus starten ( $\sigma$ -Zyklus) (Abb. 1-4). Der größte Vorteil dieses Zyklus liegt darin, dass die RNAP relativ einfach und schnell verändert werden kann, indem  $\sigma$ -Faktoren mit unterschiedlicher Promotor-Erkennungsspezifität reversibel binden und so die Transkriptionsinitiation sehr schnell an verschiedene Wachstums- und Umgebungsbedingungen anpassen können. Der oben dargestellte Mechanismus, bei dem der  $\sigma$ -Faktor sofort nach der Bildung eines stabilen RNA-DNA Hybrides die RNAP verlässt, wird als obligatorisches Modell bezeichnet (Metzger *et al.*, 1993).

Das Verlassen des  $\sigma$ -Faktors kann auch mit einem stochastischen Modell beschrieben werden. Dabei dissoziiert der  $\sigma$ -Faktor nicht sofort nach 8 bis 9 Nucleotiden von der RNAP, vielmehr nimmt die Affinität des  $\sigma$ -Faktors während der fortschreitenden RNA-Synthese kontinuierlich

ab. Die Dissoziation erfolgt in diesem Fall eher zufällig (Abb. 1-4; Shimamoto *et al.*, 1986). So liegt die Bindungskonstante für den Initiationskomplex bei ungefähr  $10^{-10}$  M und verringert sich bis auf  $10^{-6}$  M für den TEC bestehend aus RNAP und RNA-DNA Hybrid.



**Abb. 1-4: Der  $\sigma$ -Zyklus**

(A) Verschiedene  $\sigma$ -Faktoren konkurrieren um die freie ungebundene *core* RNAP (hellblau). Nachdem ein  $\sigma$ -Faktor gebunden hat, bildet sich das Holoenzym, das nun an die DNA bindet. Nach erfolgreicher Initiierung verlässt der  $\sigma$ -Faktor die RNAP. Der  $\sigma$  Zyklus ist damit abgeschlossen (Mooney *et al.*, 2005).

In (B) und (C) sind zwei unterschiedliche Wirkungsweisen des  $\sigma$ -Faktors dargestellt. Der orangefarbene Balken beschreibt die Bindungsstärke des  $\sigma$ -Faktors zur RNAP. (B) Im obligatorischen Modell bindet der  $\sigma$ -Faktor die RNAP und verlässt sie, sobald ein stabiles 8-9 bp langes RNA-DNA Hybrid gebildet wurde. (C) Nach dem stochastischen Modell besitzt der  $\sigma$ -Faktor zu Beginn der Initiierung eine hohe Affinität zur RNAP, die mit der Bildung immer längerer RNA-Transkripte kontinuierlich abnimmt (Mooney *et al.*, 2005).

Die Familie der  $\sigma$ -Faktoren in *E. coli* kann in zwei Klassen eingeteilt werden. Die erste Klasse, die  $\sigma^{70}$ -Familie, umfasst die meisten der  $\sigma$ -Faktoren, so auch den  $\sigma^{70}$ -Faktor, der die Transkription der *housekeeping* Gene kontrolliert (Borukhov und Severinov, 2002). Die zweite Gruppe umfasst die  $\sigma^{54}$ -Familie, die zur Ausbildung des Initiationskomplexes noch zusätzliche Proteinfaktoren sowie Energie benötigt, die durch ATP- oder GTP-Hydrolyse gewonnen wird (Buck *et al.*, 2000). Diese Gruppe ist zwar funktionell verwandt mit der  $\sigma^{70}$ -Familie, aber strukturell stark unterschiedlich (Paget und Helmann, 2003). Generell wird bei allen  $\sigma$ -Faktoren die DNA-Bindungsstelle durch eine Domäne blockiert, damit freie  $\sigma$ -Faktoren nicht an die entsprechenden Promotoren binden können und diese besetzen. Erst durch die Ausbildung des RNAP-Holoenzym wird diese sogenannte Autoinhibition aufgehoben (Borukhov und Severinov, 2002). Für eine kontrollierte Transkription ist es notwendig, die verschiedenen  $\sigma$ -Faktoren zu regulieren. Hierzu gibt es mehrere Möglichkeiten: Kontrolle der *de novo* Synthese sowohl auf transkriptioneller als auch translationeller Ebene, post-translationelle Modifizierung und proteolytischer Verdau. Ebenso können die  $\sigma$ -Faktoren durch die Bindung an

einen sogenannten anti- $\sigma$ -Faktor inhibiert werden. Dieser Mechanismus wurde zuerst beim Phagen T4 entdeckt. Hierbei bindet der anti- $\sigma$ -Faktor an den  $\sigma^{70}$ -Faktor und verhindert somit die Bildung des Initiationskomplexes (Helmann, 1999; Hughes und Mathee, 1998). So kann nicht nur durch das Bereitstellen unterschiedlicher  $\sigma$ -Faktoren, sondern auch durch die Limitierung bestimmter  $\sigma$ -Faktoren die Initiation an verschiedene zelluläre Bedingungen angepasst werden.

Neben den  $\sigma$ -Faktoren besitzt auch die  $\alpha$ CTD der RNAP eine wichtige regulatorische Aufgabe (Abschnitt 1.2.1). Die Regulation erfolgt über zwei verschiedene Wege. Im ersten Fall bindet  $\alpha$ CTD an eine *cis*-aktive DNA-Sequenz, die stromaufwärts des Promotors liegt und daher als *upstream promotor* (UP)-Element bezeichnet wird. Die optimale Sequenz des UP-Elementes wurde durch *in vitro* Selektion erhalten und enthält alternierend A und T (Estrem *et al.*, 1998; Estrem *et al.*, 1999). Sobald  $\alpha$ CTD an ein UP-Element bindet, wird z.B. die Aktivität des *rrnB* P1 Promotors von *E. coli* um den Faktor 30 im Vergleich zum basalen Level erhöht (Ross *et al.*, 1993). Des Weiteren interagiert  $\alpha$ CTD mit stromaufwärts der DNA gebundenen Transkriptionsaktivatoren, wie z.B. dem *catabolite activator protein* (CAP) oder dem *oxidative stress regulator protein* (OxyR) (Benoff *et al.*, 2002; Tao *et al.*, 1993; Zou *et al.*, 1992). Mutagenesestudien konnten verschiedene Bereiche auf der Oberfläche der  $\alpha$ CTD kartieren, die alle zu einem Defekt in der CAP-abhängigen Transkriptionsinitiation des *lac*-Operons führten. Die gefundenen Determinanten wurden nach Position der Aminosäure benannt, deren Variation den stärksten Effekt zeigt (Murakami *et al.*, 1996; Savery *et al.*, 1998; Zou *et al.*, 1992). Die „261-Determinante“ (Aminosäurereste 257-259 und 261) bindet an den  $\sigma^{70}$ -Faktor, die „265-Determinante“ (Aminosäurereste 265, 294, 296, 299 und 302) vermittelt eine Wechselwirkung mit dem UP-Element, die „287-Determinante“ (Aminosäurereste 285-290, 315, 317 und 318) bindet an Aktivatorproteine, wie CAP. Die Kristallstruktur von  $\alpha$ CTD gebunden an CAP und dem UP-Element konnte diese Determinanten bestätigen (Benoff *et al.*, 2002).

Einen ebenso interessanten Regulationsmechanismus stellt das SoxRS-System in *E. coli* dar, mit dessen Hilfe sich die Zelle gegen Superoxide zur Wehr setzt (Dangi *et al.*, 2004; Shah und Wolf, 2004). Im ersten Schritt bindet das konstitutiv exprimierte SoxR an die Ziel-DNA, ohne die *soxS*-Transkription zu aktivieren. Durch die Oxidation des 2Fe-2S Zentrums von SoxR wird eine Konformationsänderung induziert, die nun die *soxS*-Transkription initiiert. Das Protein SoxS bindet im Gegensatz zu CAP nicht an die „287-Determinante“ sondern an die „265-Determinante“, wodurch die  $\alpha$ CTD keine UP-Elemente mehr binden kann (Dangi *et al.*, 2004). Das exprimierte SoxS-Protein kann nun mit einer eigenen als *soxbox* bezeichneten



DNA-Sequenz interagieren, die ähnlich dem UP-Element stromaufwärts des Promotors liegt (Shah und Wolf, 2004). Dadurch werden rund 40 weitere Gene aktiviert, die den zellulären oxidativen Stress bekämpfen.

### 1.2.3. Die Elongation

Die Initiation der Transkription in Prokaryonten endet, sobald sich ein stabiles RNA-DNA Hybrid mit einer Länge von 8 bis 9 Basen gebildet hat und die Affinität des  $\sigma$ -Faktors für die Promoter-DNA und für die *core*-RNAP reduziert ist (von Hippel und Pasman, 2002). Die Elongationsgeschwindigkeit wurde im Bereich von 42 Nukleotiden (nt) pro Sekunde bis zu 90 nt/s gemessen (Condon *et al.*, 1993; Gotta *et al.*, 1991). Dabei hängt die Geschwindigkeit zum einen von Wachstumsbedingungen ab (Vogel und Jensen, 1994). Zum anderen beeinflussen Antiterminationssequenzen, wie z.B. *boxA* in ribosomalen RNA (*rrn*) Operons, die Geschwindigkeit (Vogel und Jensen, 1995). Interessanterweise besitzt das RNA-DNA Hybrid eine hohe Stabilität, im Widerspruch zu der hohen Synthesegeschwindigkeit (Nudler, 1999).

Neben der hohen Prozessivität des TEC gibt es noch zahlreiche weitere Zustände, die jeweils unterschiedliche Aufgaben erfüllen wie z.B. die Korrektur falsch eingebauter Basen (Fish und Kane, 2002). Zu den wichtigsten Zuständen des TECs gehört derjenige, bei dem die Transkription angehalten wird (Uptain *et al.*, 1997). So kann die RNAP in einem angehaltenen, wartenden oder arretierten Zustand vorkommen. Im ersten Fall wird die Transkription aufgrund von Substratmangel angehalten. Sobald genügend Substrat in Form von Nukleotiden wieder zur Verfügung steht, setzt dieser Komplex die Transkription fort.

Im zweiten Fall stoppt der TEC bei speziellen DNA-Signalen, den sogenannten Pausierstellen (*pause-sites*), um dadurch die Translation mit der Transkription zu koppeln. Daneben spielt dieser Zustand eine entscheidende Rolle bei der Termination bzw. Antitermination (Abschnitt 1.2.4 und 1.3). Die Transkription kann ebenso wie im ersten Fall ohne zusätzliche Faktoren wieder aufgenommen werden. Die *pause-sites* werden in zwei Klassen eingeteilt, in Klasse I wird eine Haarnadelstruktur in der naszierenden RNA gebildet, während Klasse II durch ein schwaches RNA-DNA Hybrid charakterisiert ist (Greive und von Hippel, 2005). Um den Einfluss der Haarnadelstruktur auf die RNAP zu untersuchen, wurde in einer Studie entweder die Klappenhelix oder der distale Teil davon deletiert (Abschnitt 1.2.1; Touloukhnov und Landick, 2003). Die daraus gewonnenen Daten weisen auf eine Interaktion der RNA-Haarnadelschleife mit der Klappendomäne hin, wodurch das aktive Zentrum der RNAP verändert wird.

Der dritte mögliche Zustand wird auch als *arrested* oder *dead-end* Komplex bezeichnet, da hier die Transkription ohne Zugabe von zusätzlichen Proteinfaktoren wie GreA und GreB, nicht mehr fortgeführt werden kann (Fish und Kane, 2002).

#### 1.2.4. Die Termination

Das Ende der Transkription wird durch Terminationssequenzen bestimmt. Sobald der TEC eine Terminationssequenz auf der DNA-Matrize erreicht, wird das RNA-DNA Hybrid destabilisiert. Dadurch dissoziiert die RNA von der DNA und leitet so die Auflösung des TECs ein. Die freigewordene RNAP steht nun für einen weiteren Transkriptionszyklus zur Verfügung. Neben der stark regulierten Transkriptionsinitiation wird auch dieser Schritt reguliert (Ciampi, 2006; Henkin, 1996; Henkin, 2000; Yarnell und Roberts, 1999). Die Termination ist entweder intrinsisch oder Rho-abhängig. Bei beiden Arten muss sich zuerst ein Wartezustand bilden (Landick, 2006). Interessanterweise stoppt nicht jede RNAP an den Terminationsstellen, vielmehr überlesen einige RNAPs das Signal und fahren mit der Transkription fort. Aufgrund dieser Tatsache werden die Terminationsstellen durch zwei Größen charakterisiert: (I) durch die Effizienz des Signals, bestimmt durch die Anzahl der angehaltenen RNAPs und (II) durch die Verweildauer der RNAPs am Terminationssignal, die im Bereich von wenigen Sekunden bis hin zu einer Minute liegen kann (Neuman *et al.*, 2003).

Bei der intrinsischen Termination werden keine zusätzlichen Proteine benötigt. Das Signal hierzu ist mit der *pause-site* vergleichbar (Abschnitt 1.2.3) und besteht aus einer palindromischen Sequenz, die die Ausbildung einer RNA-Haarnadelschleife induziert. Es folgt eine  $\sim 7$  bp lange uracilreiche Sequenz, die nur ein schwaches RNA-DNA Hybrid bilden kann. Als erstes wird durch den intrinsischen Terminator ein Warten der RNAP induziert, wodurch eine vollständige Faltung der Terminationshaarnadelschleife erfolgen kann. Die Faltung der Haarnadelschleife bewirkt, wie auch bei den *pause-sites*, eine für diese Art der Termination wichtige konformationelle Änderung der RNAP.

Derzeit existieren drei mögliche Mechanismen, um das Verhalten des Terminationskomplexes zu beschreiben. Das allosterische Modell schlägt eine Interaktion der Terminationshaarnadelschleife mit der Klappendomäne sowie einer  $Zn^{2+}$ -Bindungsdomäne, die in der  $\beta'$  Unter-einheit liegt, vor. Die dadurch induzierten Konformationsänderungen bewirken ein Aufschmelzen des RNA-DNA Hybrids, wodurch die Transkriptionsblase zusammenbricht. Für dieses Modell sprechen Mutationsexperimente, bei denen die Klappendomäne und

Zn<sup>2+</sup>-Bindungsdomäne in der RNAP deletiert wurden. Daraus resultierte eine geringere Terminationsrate (Toulokxonov und Landick, 2003; Toulokxonov *et al.*, 2001).

Bei dem sog. Hybridaufschmelz-Modell wird die notwendige konformationelle Änderung des TECs ausschließlich durch das Aufschmelzen des RNA-DNA Hybrides hervorgerufen (Gusarov und Nudler, 1999; Komissarova *et al.*, 2002).

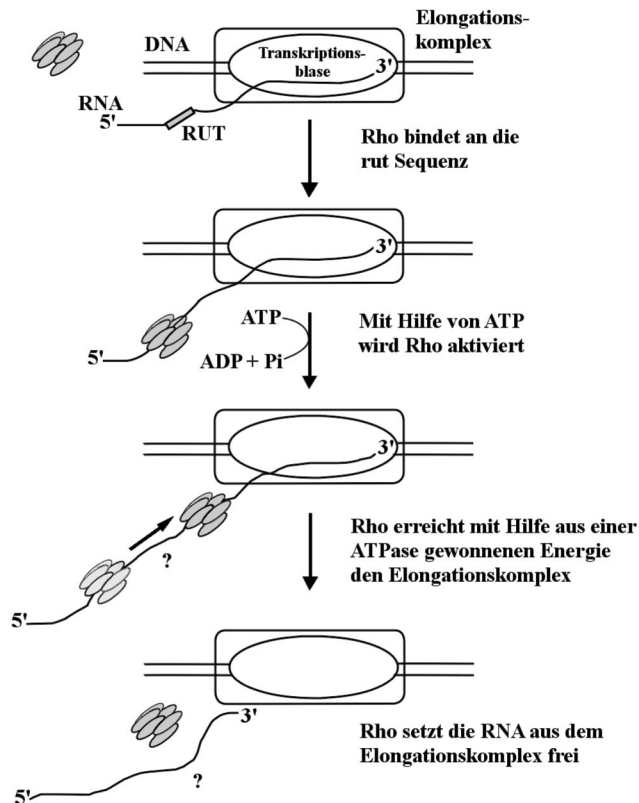
Ebenso verringert das Vernetzen der stromabwärts gelegenen DNA die Terminationsrate. Da in diesem Fall die Vorwärtsbewegung der RNAP behindert wird, spricht man vom Vorwärtstranslokationsmodell. Bei diesem Modell bewegt sich der Terminationskomplex vorwärts, ohne Nukleotide einzubauen. Dadurch wird das RNA-DNA Hybrid verkürzt und so stark destabilisiert, dass eine Termination erfolgt (Yarnell und Roberts, 1999).

Das Signal für die Rho-abhängige Termination ist aus zwei Teilen aufgebaut. Der erste Teil besteht aus einer *Rho utilization (rut)* Sequenz, die entweder aus einer 30 nt langen Sequenz, die 35 % - 65 % Cytosin enthält, oder aus einer 6 nt umfassenden Haarnadelstruktur besteht (Schneider *et al.*, 1993). Der zweite Teil enthält die Terminationsregion, die 60 bis 90 nt stromabwärts der *rut* Sequenz liegt. Die Termination erfolgt innerhalb dieser Zone, allerdings an keiner exakt definierbaren Position (Banerjee *et al.*, 2006).

Für die Rho-abhängige Termination ist vor allem das Rho-Protein notwendig. Dieses Protein ist in allen Bakterien essentiell für das Überleben. Es ist aus einer hexameren Ringstruktur gleicher, 46,8 kDa großer Untereinheiten aufgebaut. Die Aufgabe des Rho-Proteins besteht darin, als RNA/DNA-Helikase das RNA-DNA Hybrid im TEC aufzuschmelzen, um so die Transkription abubrechen. Die hierfür nötige Energie wird durch ATP-Hydrolyse gewonnen. Die aminoternale Domäne eines Monomers enthält die primäre RNA-Bindungsstelle. Die carboxyterminale Domäne enthält drei besondere Strukturmerkmale, eine P-Schleife, die die ATP-Bindung und Hydrolyse durchführt, sowie einer R-Schleife und einer Q-Schleife, die zusammen die sekundäre RNA-Bindungsstelle bilden und im Inneren des durch die sechs Monomere gebildeten Ringes liegen (Skordalakes und Berger, 2003).

Der Mechanismus der Rho-abhängigen Termination erfolgt in mehreren Schritten. Zunächst bindet das Rho-Protein an die *rut* Sequenz (Abb. 1-5). Aufgrund der Orientierung der sechs primären RNA-Bindungsstellen wird das 3' Ende der RNA in die Nähe der sekundären Bindungsstelle gebracht. Die Bindung der RNA induziert die Ausbildung einer geschlossenen stabilen Ringstruktur. Das Rho-Protein kann sich nun unter ATP-Hydrolyse entlang der RNA zum TEC bewegen. Der genaue Mechanismus ist bis jetzt noch weitgehend unverstanden. Im letzten Schritt, bei dem das RNA-DNA Hybrid aufgeschmolzen wird, erfolgt der Abbruch der

Transkription, wobei ein ähnlicher Mechanismus wie für die intrinsische Termination vorgeschlagen wird (Greive und von Hippel, 2005; Skordalakes und Berger, 2003).



#### Abb. 1-5: Überblick über die einzelnen Schritte der Rho-abhängigen Termination

Der Rho-Faktor bindet im ersten Schritt an die rut Sequenz. Unter ATP-Hydrolyse bewegt sich Rho zum TEC, um dort die Transkription zu terminieren. Die Translokation des Rho-Faktors, die Wechselwirkung mit dem Elongationskomplex und das Entwinden des RNA-DNA Hybrides sind bisher noch unverstanden (Banerjee *et al.*, 2006).

### 1.3. Antitermination

Die Termination der Transkription erfolgt an den wenigen vorhandenen Terminationsstellen nur mit geringer Effizienz. Deswegen besteht eine Möglichkeit zur Regulation der Transkription darin, die Effizienz an einem Terminationsignal positiv oder negativ zu beeinflussen (Weisberg und Gottesman, 1999). Negative Regulation des Terminationsignales führt zur Fortführung der Transkription durch Überlesen des Terminationsignales. Man unterscheidet zwei unterschiedliche Arten, erstens die Attenuation, die an *his* und *trp* Operons untersucht wurde (Henkin, 1996), zweitens die prozessive Antitermination. Bei diesem Prozess werden nicht nur nachfolgende Terminationsignale überlesen, sondern zusätzlich wird die Sensitivität der RNAP gegenüber Wartestellen im Genom verringert (Weisberg und Gottesman, 1999). Die Antitermination des Phagen  $\lambda$  stellt den Prototyp dieses Mechanismus dar, sie wird aber auch im Genom von *E. coli* gefunden. Vor allem die Transkription der ribosomalen RNAs (*rrn*) wird häufig mit Hilfe von Antitermination reguliert.

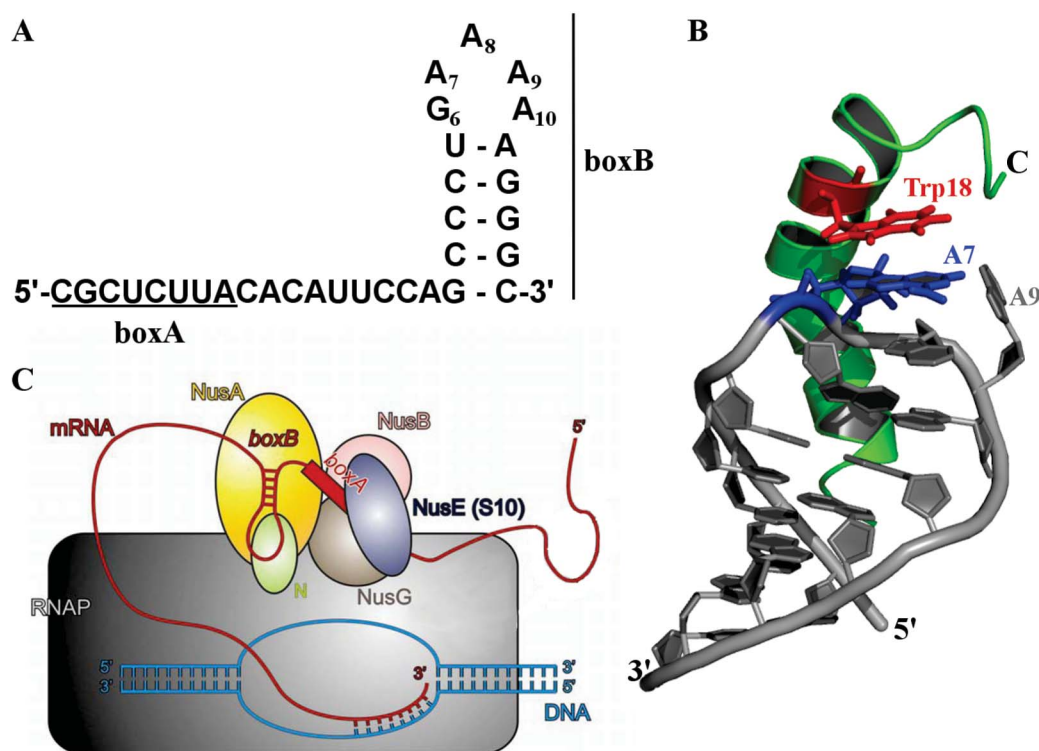
### 1.3.1. Der Antiterminationsmechanismus des Phagen $\lambda$

Das Signal für die Antitermination des Phagen  $\lambda$  ist eine RNA-Sequenz, die als *N utilization site* (*nut*) bezeichnet wird (Friedman und Baron, 1974; Nudler und Gottesman, 2002). Diese Sequenz ist zweimal im Phagengenom enthalten (Abb. 1-1), zum einen ca. 50 bp stromabwärts des  $p_L$  Promotors (*nutL*), zum anderen ca. 250 bp stromabwärts des  $p_R$  Promotors (*nutR*) (Drahos *et al.*, 1982; Rosenberg *et al.*, 1978). Beide Terminationssequenzen bestehen aus einem *boxA* und *boxB* Motiv, die durch eine *Spacer* Sequenz getrennt sind (Abb. 1-6 A) (Hasan und Szybalski, 1986). Die *boxA* Sequenz ist bei der *nutL* und bei der *nutR* RNA gleich, wohingegen in der *nutL boxB* RNA an Position 9 ein Guanin anstelle von Adenin zu finden ist (Abb. 1-6). *boxA* umfasst 8 nt und wird ebenso in der ribosomalen RNA gefunden (Abschnitt 1.3.2). *boxB* ist ein unterbrochenes Palindrom mit einer Stammregion und einer aus fünf Basen bestehenden Schleife (Abb. 1-6A), die jedoch nicht in dieser Form in der ribosomalen RNA zu finden ist (Abschnitt 1.3.2).

Für die Antitermination des Phagen  $\lambda$  ist vor allem das Protein  $\lambda N$  wichtig (Greenblatt *et al.*, 1980). Die Regulation von  $\lambda N$  und somit indirekt die Regulation der Antitermination kann auf drei unterschiedliche Arten erfolgen. Erstens, es kann durch die Proteine Cro und CI die Synthese von  $\lambda N$  gehemmt werden, indem der Promotor  $p_L$  blockiert wird (Court *et al.*, 2007). Zweitens, es kann die Aktivität des unstrukturierten  $\lambda N$  durch den Abbau mit Hilfe verschiedener Proteasen, z.B. Lon, reguliert werden (Court *et al.*, 2007; Maurizi, 1987). Drittens, es besteht die Möglichkeit, dass sich  $\lambda N$  mit Hilfe der RNase III selbst reguliert (Court *et al.*, 2007; Wilson *et al.*, 2002). Durch die Bildung des Antiterminationskomplexes an der *nutL* Position wird die für den Start der Expression von  $\lambda N$  wichtige Shine-Dalgarno-Sequenz (SD) verdeckt. Erst durch spezifisches Schneiden der RNA zwischen der *nutL* und der  $\lambda N$  SD-Sequenz wird letztere wieder für das Ribosom zugänglich, so dass  $\lambda N$  exprimiert werden kann.

Das 107 Aminosäure lange intrinsisch unstrukturierte  $\lambda N$  weist drei unterschiedliche Sequenzbereiche auf: Aminosäuren 1-22 binden die *boxB*, 34-47 binden den carboxyterminalen Teil von *E. coli* NusA (Abschnitt 1.4) und 73-107 kontaktieren die RNAP (Greenblatt *et al.*, 1998; Mogridge *et al.*, 1998). Der aminoterminalen Teil (Aminosäuren 1-22) gehört zur Familie der Arginin-reichen Motive (ARM) und bindet mit einer Affinität von  $\sim 10^{-9}$  M an *boxB* (Cilley und Williamson, 1997; Tan und Frankel, 1995; Van Gilst und von Hippel, 1997). In der mit NMR-Spektroskopie gelösten Struktur bindet das basische ARM des  $\lambda N$  Proteins in Form einer gebogenen  $\alpha$  Helix an das 5'-Ende der Stammschleifenstruktur von *boxB* (Abb.

1-6B; Legault *et al.*, 1998; Scharpf *et al.*, 2000). Die gefaltete *boxB* zeigt starke strukturelle Ähnlichkeiten mit einem GNRA-Tetraloop. GNRA-Schleifen werden durch nicht reguläre Watson-Crick-Basenpaare zwischen dem ersten und vierten Nukleotid der entsprechenden Schleife gebildet und weisen eine hohe thermische Stabilität auf (Jucker *et al.*, 1996). Abweichend zur Struktur des GNRA-Loops ist bei *boxB* die vierte Base der Schleife (Adenin 9) nach außen gedreht (Abb. 1-6 B). Der Komplex aus  $\lambda$  N und der *boxB* RNA wird durch  $\pi$ - $\pi$  Wechselwirkung von Adenin 7 (A7) mit dem aromatischen Ring von Trp18 (Abb. 1-6 B) sowie durch zusätzliche hydrophobe und elektrostatische Wechselwirkungen im Stammbereich stabilisiert (Legault *et al.*, 1998; Scharpf *et al.*, 2000; Xia *et al.*, 2005).



**Abb. 1-6: Wichtige Elemente des Antiterminationskomplexes**

(A) Die *nutR* RNA besteht aus *boxA* und *boxB*. *boxB* formt einen GNRA Tetraloop, wobei das Adenin 9 aus dem Loop geklappt ist (Legault *et al.*, 1998). (B) Das ARM Motiv des  $\lambda$  N Peptides bindet in Form einer geknickten Helix (grün) an den 5' Stamm der *nutR boxB*, wobei Trp 18 (rot) über  $\pi$ - $\pi$  Wechselwirkung Adenin 7 (blau) bindet (PDB 1QFQ; Legault *et al.*, 1998; Scharpf *et al.*, 2000). (C) Durch Bindung weiterer Antiterminationsfaktoren, wie NusA, NusB, NusE oder NusG wird der Komplex stabilisiert (Bonin, 2004; Mogridge *et al.*, 1998).

Verschiedene *in vitro* und *in vivo* Studien zeigten, dass die Wechselwirkung zwischen Trp18 und A7 sowohl für die Stabilität des Komplexes als auch für die Effizienz der Antitermination wichtig ist (Franklin, 1993; Su *et al.*, 1997). Die Bedeutung der Trp18/A7-Wechselwirkung wurde durch *in vitro* Proteinselektion festgestellt (Xia *et al.*, 2003b). Hierbei wurde unabhängig von der Affinität des  $\lambda$  N Proteins für die RNA eine hohe biologische Aktivität nur bei

einer Interaktion zwischen Trp18 und A7 erreicht. Die Positionen 14 und 15, die beide eine Helix Windung von Trp18 entfernt sind, spielen für die Affinität eine entscheidende Rolle, wie Mutationen an diesen Positionen nahe legen (Xia *et al.*, 2003a; Xia *et al.*, 2003b; Xia *et al.*, 2005). Mittels Femtosekundenspektroskopie wurde ein Gleichgewicht zwischen einem offenen und einem geschlossenen Zustand des carboxyterminalen Teils der Helix (Aminosäuren 12-22) im  $\lambda$  N/*boxB* Komplex gefunden (Xia *et al.*, 2003a). Im offenen Zustand wird keine Wechselwirkung zwischen Trp18 und A7 beobachtet, im geschlossenen schon. Das Gleichgewicht wird von den Aminosäurenresten an den Positionen 14 und 15 beeinflusst (Xia *et al.*, 2005). In den Experimenten löst die Bindung von  $\lambda$  N an die RNA alleine noch keine Antitermination aus. Erst die sehr spezifische Wechselwirkung zwischen Trp18 und A7 bewirkt eine Umorientierung und Veränderung der RNAP zu einem terminationsresistenten Komplex, wobei der genaue Wirkmechanismus dieser Interaktion noch nicht verstanden ist. Obwohl  $\lambda$  N *in vitro* an proximal zum Promotor gelegenen Terminationsstellen eine Antitermination hervorrufen kann, werden für einen prozessiven Antiterminationskomplex, der weit vom Promotor entfernte Terminationsstellen überlesen kann, weitere Wirtsfaktoren benötigt (Abb. 1-6 C; Rees *et al.*, 1996). Diese werden als *N utilization substance* (Nus) bezeichnet. Eines der wichtigsten Proteine ist der Elongationsfaktor NusA (Abschnitt 1.4; Whalen *et al.*, 1988), der direkt mit  $\lambda$  N wechselwirkt (Bonin *et al.*, 2004). Des Weiteren sind die an *boxA* bindenden Protein NusE und NusB an diesem Vorgang beteiligt (Altieri *et al.*, 2000; Greive *et al.*, 2005). NusE wird auch als S10 bezeichnet, da es sowohl im Antiterminationskomplex als auch im Ribosom vorhanden ist (Friedman *et al.*, 1981). Alle zusammen liegen mit NusG in einen stabilen prozessiven Antiterminationskomplex vor (Abb. 1-6 C; Ciampi, 2006; Li *et al.*, 1993; Mason und Greenblatt, 1991). Die einzelnen Proteine bilden ein Gerüst, das die RNAP so verändert, dass es Terminationsstellen nicht mehr erkennt. Die *nut* Sequenz hat die Aufgabe, die einzelnen Faktoren zu binden und die lokale Konzentration der Antiterminationsfaktoren in der Nähe der RNAP zu erhöhen (Das *et al.*, 1996; Greive und von Hippel, 2005; Mason und Greenblatt, 1991). Diese vielfältigen Interaktionen stabilisieren den Transkriptionskomplex, was zu einem kürzeren Verweilen an Terminationsstellen und somit zu einer geringeren Terminationseffizienz führt (Gusarov und Nudler, 2001; Nudler und Gottesman, 2002). Denkbar wäre auch eine Stabilisierung des RNA-DNA Hybrids durch eine Veränderung des RNA-Austrittskanals im Antiterminationskomplex (Greive und von Hippel, 2005). Dies hätte ebenfalls eine geringere Terminationseffizienz zur Folge.

### 1.3.2. Die Antitermination an *rrn* Operons

In *E. coli* gibt es sieben verschiedene *rrn* Operons, die als *rrnA-E* und *rrnG-H* bezeichnet werden. Obwohl keinem der sieben Operons eine bestimmte Aufgabe zugeordnet werden konnte, sind alle sieben Operons für eine schnelle Anpassung an Wachstumsbedingungen notwendig (Condon *et al.*, 1992). Der Grundaufbau aller sieben Operons ist relativ einheitlich. Den Anfang bildet eine Leitsequenz gefolgt von den 16S-rRNA-Genen. Daran schließt sich eine Trennsequenz an, die je nach Operon unterschiedliche tRNA Gene enthält. So enthalten die *rrns B, C, E* und *G* die tRNA<sup>Glu-2</sup>, und die *rrns A, D* und *H* die tRNA<sup>Ala-1B</sup> und tRNA<sup>Ile-1</sup>. Die Trennsequenz separiert die 16S-Gene von den 23S-Genen. Vor dem Terminationssignal liegen die 5S-rRNA Gene. Die Leitsequenz enthält unterschiedliche regulatorische Elemente: zwei  $\sigma^{70}$ -Promotoren ( $p_1$  und  $p_2$ ), zwei unterschiedliche Aktivatorsequenzen (Abschnitt 1.2.2) sowie ein Antiterminationssignal (Condon *et al.*, 1995; Morgan, 1986).

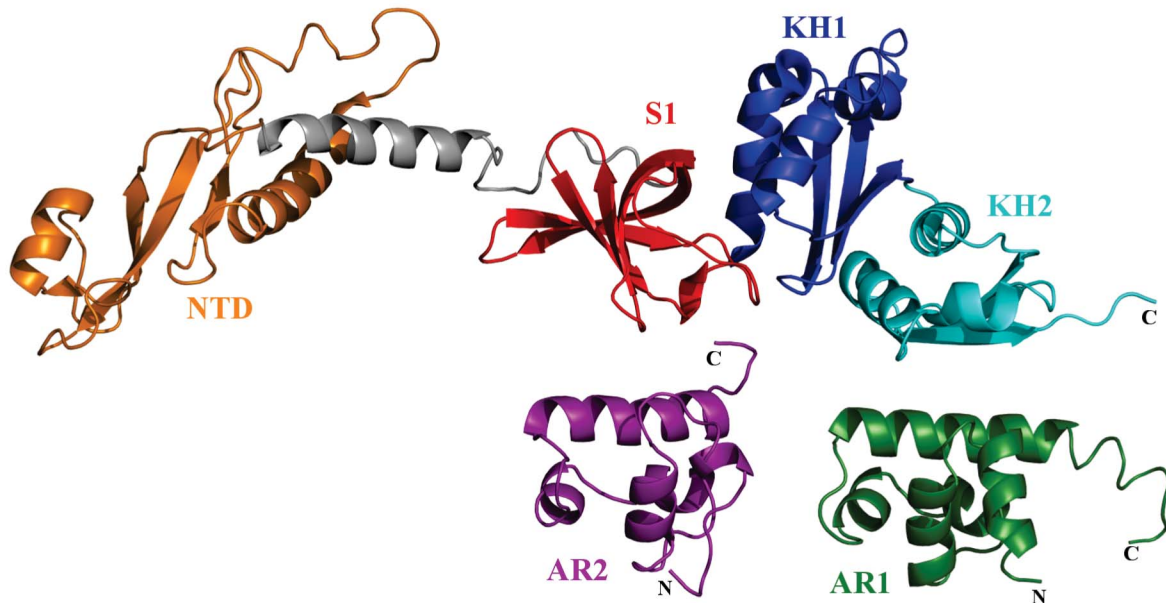
Das Antiterminationssignal wurde aufgrund der hohen Sequenzhomologie zu  $\lambda$  *nutL* und *nutR* gefunden. *In vivo* Studien bestätigten die Relevanz dieser Sequenz für die Antitermination (Li *et al.*, 1984). Sowohl in der *rrn* Leitsequenz als auch in der Trennsequenz konnte jeweils eine *boxA*, *boxB* und *boxC* Sequenz gefunden werden. Allerdings liegen *boxB* und *boxA* im Vergleich zum Phagen  $\lambda$  in umgekehrter Reihenfolge vor (Abschnitt 1.3.1). *rrn boxB* besteht aus einer Stamm-Schleife-Struktur. Anders als bei  $\lambda$  ist diese *boxB* Sequenz nicht konserviert und wird für eine Antitermination nicht unbedingt benötigt (Gourse *et al.*, 1986), Mutationen in der stark konservierten *boxA* führen dagegen zum Verlust der Antitermination (Heinrich *et al.*, 1995). Die starke Homologie zum Phagen  $\lambda$  System legt nahe, dass im Falle der *rrn*-Antitermination ähnliche Proteinfaktoren beteiligt sind. Zunächst wurde die Bindung von NusB sowie NusE an *boxA* gezeigt (Greive *et al.*, 2005; Squires *et al.*, 1993). Weitere Studien konnten eine Beteiligung von NusA und NusG belegen (Liu und Hanna, 1995; Mason *et al.*, 1992; Squires *et al.*, 1993; Ward und Gottesman, 1981).

Ein entscheidender Unterschied zum  $\lambda$  N System besteht darin, dass das *rrn*-Antiterminationssystem nur gegen Rho-abhängige Termination aktiv ist, während das  $\lambda$  System sowohl gegenüber Rho-abhängiger als auch intrinsischer Termination aktiv ist (Albrechtson *et al.*, 1990; Condon *et al.*, 1995). Die Aufgabe der *rrn*-Antitermination besteht in der Regulation der Transkriptionsgeschwindigkeit. Durch eine langsamere Geschwindigkeit bleibt genügend Zeit für eine korrekte Faltung der rRNA (Squires *et al.*, 1993).



## 1.4. Der Elongationsfaktor NusA

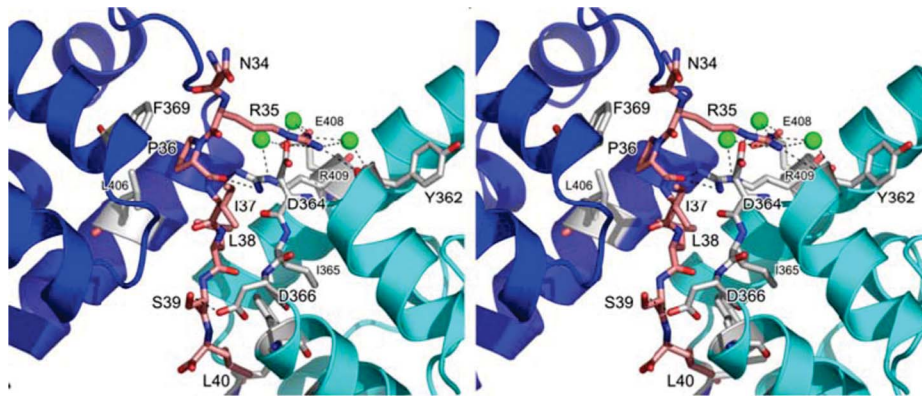
Der Elongationsfaktor NusA wurde zuerst durch genetische Studien als Bestandteil des  $\lambda$  N Antiterminationskomplexes identifiziert (Friedman und Baron, 1974). NusA ist ein essentieller, multifunktionaler Transkriptionselongationsfaktor, der sowohl in Eubakterien als auch in Archaea zu finden ist (Nudler und Gottesman, 2002). Das 495 Aminosäuren umfassende Protein besteht aus sechs Domänen. Das Zentrum wird durch drei RNA-Bindungsdomänen gebildet, eine S1-Domäne, homolog zu einem RNA-Bindungsmotiv des ribosomalen Proteins S1, sowie zwei KH-Domänen (KH1 und KH2), die homolog zum mRNA Bindungsprotein K sind. In den mit Röntgenkristallographie bestimmten homologen NusA Strukturen aus *Thermotoga maritima* (*T. maritima*) und *Mycobacterium tuberculosis* (*M. tuberculosis*) liegen die RNA-Bindungsdomänen in elongierter Form vor und sind starr miteinander verbunden (Abb. 1-7; Gopal *et al.*, 2001; Worbs *et al.*, 2001). Die Hauptaufgabe von NusA ist die Wechselwirkung mit der naszenten RNA. In der Kristallstruktur von *M. tuberculosis* wird die naszierende *rrn* *BoxC* in gestreckter Form von den beiden KH Domänen gebunden (Arnvig *et al.*, 2004). Die aminoterminal Domäne (NTD), die möglicherweise mit RNAP interagiert, ist über eine flexible Verbindungshelix mit dem Zentrum von NusA verbunden (Gopal *et al.*, 2001; Worbs *et al.*, 2001). Im Vergleich zu *M. tuberculosis* und *T. maritima* besitzt *E. coli* zwei weitere carboxyterminale Domänen, die als *acidic repeat* 1 und *acidic repeat* 2 bezeichnet werden (AR1 und AR2) (Abb. 1-7; Eisenmann *et al.*, 2005). Die zelluläre Funktion dieser stark sauren Domänen wurde bisher noch nicht vollständig aufgeklärt. Eine der derzeit bekannten zellulären Aufgaben von AR2 besteht darin, mit den drei RNA-Bindungsdomänen zu interagieren. Dadurch wird eine Bindung von freiem NusA an RNA verhindert. Erst durch die Interaktion von AR2 mit  $\alpha$ CTD wird diese Selbstblockade, die als Autoinhibition bezeichnet wird, aufgehoben (Mah *et al.*, 2000).



**Abb. 1-7: Struktur von *T. maritima* NusA und den beiden *E. coli* Domänen AR1 und AR2**

Die aminoterminal Domäne (NTD) (orange) von *T. maritima* NusA ist über eine Helix (grau) mit den drei RNA-Bindungsdomänen S1 (rot), KH1 (blau) und KH2 (cyan) verbunden (PDB 1HH2)(Worbs *et al.*, 2001). Die in *E. coli* zusätzlich vorhandenen Domänen AR1 und AR2 sind in grün und violett dargestellt (PDB 1WCL und 1WCH; Eisenmann *et al.*, 2005). Die AR2 Domäne verhindert die Bindung von freiem *E. coli* NusA an RNA. Dies wird als Autoinhibition bezeichnet (Mah *et al.*, 1999).

Die zelluläre Funktion von AR1 ist bisher noch unbekannt. Im  $\lambda$  N Antiterminationskomplex interagiert AR1 mit  $\lambda$  N (Bonin *et al.*, 2004; Greenblatt *et al.*, 1998; Gusarov und Nudler, 2001; Mogridge *et al.*, 1998). In der Kristallstruktur des Komplexes aus  $\lambda$  N und NusA binden die Aminosäuren 34-40 des  $\lambda$  N Proteins in gestreckter Konformation an zwei AR1 Moleküle (AR1<sup>A</sup> und AR1<sup>B</sup>) (Abb. 1-8; Bonin *et al.*, 2004). Da AR2 im Kristall durch Proteasen abgebaut wurde, kann eine Interaktion von  $\lambda$  N und AR2 nicht ausgeschlossen werden. Mutationsstudien zeigten AR1<sup>B</sup> als die biologisch aktive Form. Hierfür sprechen auch die größere Bindungsfläche des  $\lambda$  N/AR1<sup>B</sup> Komplexes sowie das elektrostatische Potential an der entsprechenden Kontaktfläche. Vor allem hydrophobe Wechselwirkungen der Leucine 38 und 40 von  $\lambda$  N stabilisieren die gestreckte Konformation. Die Seitenkette von Leu-40 bildet hydrophobe Wechselwirkungen mit Leu-398 und Val-372 von AR1 aus. Darüber hinaus stapelt die Seitenkette mit dem aromatischen Ring von Phe-369 und füllt damit eine hydrophobe Tasche (Abb. 1-8; Bonin *et al.*, 2004). In der Kristallstruktur liegen außerdem drei Wassermoleküle vor, die zur Stabilisierung des  $\lambda$  N-AR1<sup>B</sup> Komplexes beitragen (Abb. 1-8).

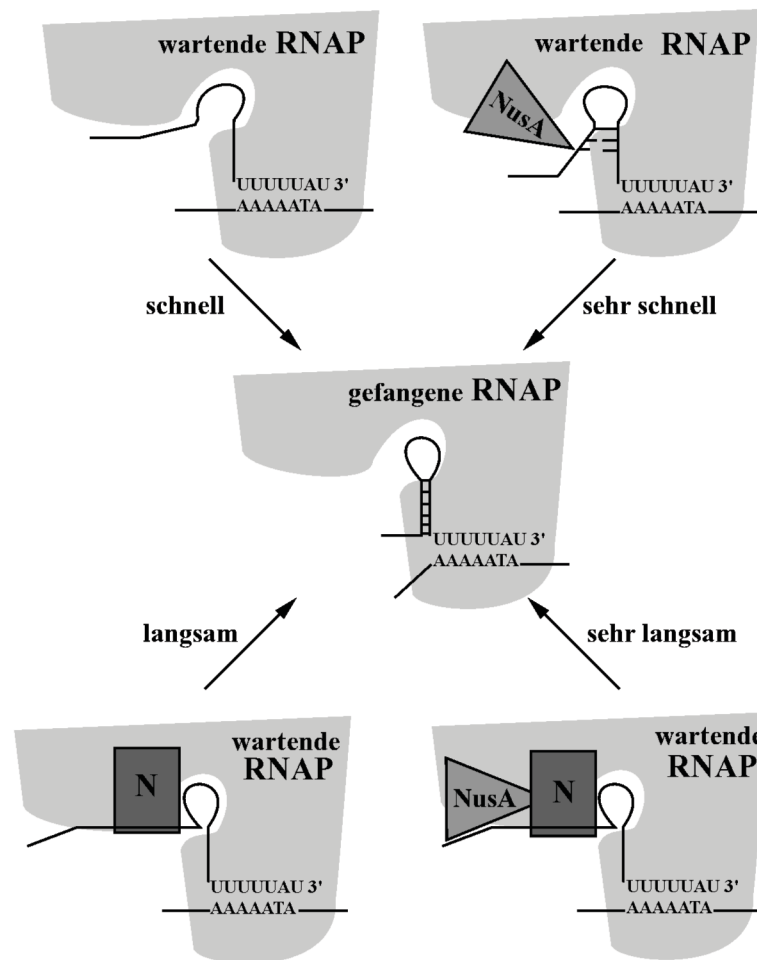


**Abb. 1-8: Stereodarstellung des  $\lambda$  N(34-40) NusA-AR1 Komplex**

Es sind die AR1<sup>A</sup> und AR1<sup>B</sup> Moleküle in blau und cyan dargestellt. Aminosäurereste, die mit  $\lambda$  N interagieren sind als Stabmodell dargestellt (Kohlenstoff grau, Sauerstoff rot und Stickstoff blau). Das  $\lambda$  N Protein ist ebenso als Stabmodell gezeichnet, wobei der Kohlenstoff rosa gefärbt ist. Die an der Interaktion beteiligten Wassermoleküle (grün) stabilisieren die Bindung. AR1<sup>A</sup> stellt ein Kristallartefakt dar, während AR1<sup>B</sup> an der biologischen Funktion beteiligt ist (Bonin *et al.*, 2004; Bonin, 2004).

Die Multidomänenstruktur von NusA spiegelt sich auch in der Multifunktionalität wider (Borukhov *et al.*, 2005). So fördert NusA primär die Termination, sobald die *nut* oder *rrn* Antiterminationssignale vorhanden sind verstärkt es die Antitermination (Gusarov und Nudler, 2001). Mit Hilfe von Vernetzungsexperimenten wurde folgende vorläufige Wirkungsweise von NusA vorgeschlagen (Abb. 1-9): Die für die Termination wichtige Haarnadelschleife hat zur Faltung rund 2 s Zeit (Gusarov und Nudler, 2001), bevor die RNAP die Terminationsstelle verlässt, und so die Transkription fortgesetzt wird. Die Transkription pausiert, so dass genügend Zeit zur Ausbildung der transkribierten Haarnadelschleife bleibt. Der bestimmende Schritt der Termination ist also die Ausbildung der Haarnadelschleife, die zu einem geschwächten RNA-DNA Hybrid führt (Abschnitt 1.2.4). Für die Entstehung der Haarnadelschleife sind zwei Punkte wichtig: erstens muss das RNA-DNA Hybrid partiell aufgeschmolzen werden, zweitens müssen Kontakte der naszierenden RNA mit der stromaufwärts gelegenen Bindungsstelle der RNAP gelockert werden (Abb. 1-2). Diese stromaufwärts gelegene Bindungsstelle besteht aus dem aminoterminalen Bereich der  $\beta'$  Untereinheit (Aminosäurereste 1 bis 102), dem carboxyterminalen Bereich der  $\beta$  Untereinheit (Aminosäurereste 1304 bis 1342) und aus Teilen der  $\alpha$ CTD (Korzheva *et al.*, 2000; Liu *et al.*, 1996; Nudler, 1999). Vernetzungsexperimente lassen auf eine Schwächung der Interaktion der naszierenden RNA mit der RNAP schließen. Dies führt zu einer schnelleren Bildung der Haarnadelschleife (Abb. 1-9) und somit zu einer höheren Terminationseffizienz. Ist aber das  $\lambda$  N Protein vorhanden, wird ein anderer Weg eingeschlagen (Abb. 1-9).  $\lambda$  N verhindert die Ausbildung der Terminationsschleife und verändert zudem NusA. NusA fördert nun nicht mehr die Bildung der Haarnadelschleife sondern bindet den 5' Stamm der Haarnadelschleife. Durch die Vernetzungsexperimente wurde diese Haarnadelschleifen-Bindung im Bereich der Aminosäurereste 167 bis

204 von NusA nachgewiesen, die alle in der S1-Bindungsdomäne liegen (Gusarov und Nudler, 2001). Da sich nun die notwendige Terminationsschleife nicht ausbilden kann, erfolgt keine Termination (Gusarov und Nudler, 2001). Bei diesem Mechanismus stellt sich die Frage, wie ein NusA Molekül den 5' Stamm der Haarnadelschleife binden und dennoch gleichzeitig an der Interaktion mit der *nut* oder *rrn* RNA beteiligt sein kann. Da Western blot Analysen eine Bindung von zwei NusA Molekülen an die RNAP zeigen (Horwitz *et al.*, 1987), könnte ein Molekül die *nut* oder *rrn* RNA binden, während das zweite Molekül den 5' Stamm der Terminationshaarnadelschleife binden würde (Gusarov und Nudler, 2001).



**Abb. 1-9: Das Model für die Wirkungsweise von NusA nach** (Gusarov und Nudler, 2001)

Es sind zwei Zustände der RNAP (grau) zu sehen, ein wartender Zustand, der die Transkription fortführen kann, und ein gefangener Zustand, der zum Abbruch der Transkription führt. Im ersten Fall wird durch NusA die Transkription erhöht, im zweiten Fall wird die Terminationseffizienz mit Hilfe von  $\lambda$  N herabgesetzt. Durch die Anwesenheit von NusA wird die Effizienz der Antitermination deutlich erhöht, da sich der gefangene Zustand nur sehr langsam ausbildet.

## 2. Ziele

Die Antitermination wurde zuerst im Phagen  $\lambda$  entdeckt und ist ein zentrales Element in der Regulation der Transkription. Dennoch ist der exakte Mechanismus bisher noch nicht bekannt. Vor allem die Rolle des hierzu notwendigen Elongationsfaktors NusA ist noch weitgehend ungeklärt. Es sind zwar homologe Strukturen bekannt, aber nicht die Struktur von NusA aus *E. coli*. Ebenso sind noch nicht alle zellulären Funktionen im Detail verstanden. Ein zentrales Ziel dieser Arbeit war daher, die Wechselwirkungen des essentiellen Elongationsfaktors NusA mit verschiedenen Bindungspartnern im Antiterminationskomplex auf atomarer Ebene zu untersuchen.

In der Kristallstruktur bindet NusA-AR1 an  $\lambda$  N, jedoch fehlten sowohl AR2 als auch Teile von  $\lambda$  N, die aufgrund biochemischer Daten zum Bindungsmotiv gehören. Darauf basierend wurde eine Bindung dieses Teils an die zweite Domäne AR2 vermutet. Die vorliegende Arbeit sollte unter anderem klären, ob der fehlende Teil von  $\lambda$  N an AR2 oder AR1 bindet und ob die Bindung von NusA an  $\lambda$  N die Interaktion zwischen  $\lambda$  N und *boxB* beeinflusst.

Zur Regulation der Elongation ist sowohl NusA als auch die  $\alpha$ CTD der RNAP von zentraler Bedeutung. Die aufgrund von Literaturdaten (Mah *et al.*, 2000) vermutete Interaktion zwischen NusA-AR2 und  $\alpha$ CTD sollte auf atomarer Ebene studiert und die daraus gewonnenen Erkenntnisse zur Klärung des regulatorischen Einflusses von NusA auf die Elongation herangezogen werden.

Da beim  $\lambda$  System NusA vermutlich an den 3' Stamm der *boxB* bindet (Zhou *et al.*, 2001), die *boxB* bei der *rrn* Antitermination aber keine Rolle spielt, sollte mit Hilfe von Fluoreszenzmessungen die genauen Bindungssequenzen für beide Fälle quantitativ untersucht werden.

### 3. Zusammenfassung und Diskussion der Ergebnisse

#### 3.1. Charakterisierung von Proteinkomplexen

Um die zahlreichen Bindungspartner des Elongationsfaktors NusA auf struktureller Ebene untersuchen zu können, wurde Kernspinresonanzspektroskopie (*nuclear magnetic resonance*, NMR) verwendet. Im Gegensatz zur Röntgenkristallografie kann in der NMR-Spektroskopie in Lösung gearbeitet werden. Vor allem transiente Komplexe, wie sie bei der Antitermination eine wichtige Rolle spielen, sind oft nicht zu kristallisieren. Ein Nachteil der NMR-Spektroskopie liegt in der Größenlimitierung. Proteine oder Proteinkomplexe mit einer molekularen Masse von mehr als 30 kDa, sofern es sich um eine einzelne Polypeptidkette handelt, sind nur sehr schwer zugänglich, da zunehmende transversale Relaxation mit steigender Molekülmasse zu nicht mehr ausreichendem Magnetisierungstransfer führt (Clare und Gronenborn, 1998). Neben dem dadurch bedingten Signalverlust führt auch die Signalüberlagerung zu Problemen. Daher wurden in dieser Arbeit gezielt die Gene einzelner Proteindomänen kloniert, um das Größenproblem zu umgehen.

Um strukturelle Änderungen in Komplexen, die durch Bindung von anderen Molekülen verursacht werden, genauer zu charakterisieren, ist die Beobachtung von Änderungen der NMR-chemischen-Verschiebung (*chemical shift perturbation*; CSP) eine hervorragend geeignete Methode (Bernd Meyer, 2003; Clarkson und Campbell, 2003; Zuiderweg, 2002). Bei der Bindung eines Liganden an ein Rezeptorprotein werden die chemischen Verschiebungen der Signale beider Partner, die des Liganden und die des Proteins, verändert. Da die Resonanzen von Kernen, die sich in unmittelbarer Nähe zur Bindungsstelle befinden, die größten Änderungen der chemischen Verschiebung zeigen, lässt sich die Bindungsfläche auf diese Weise sehr einfach kartieren. Um eine eindeutige Zuordnung der Protonenresonanzen zu ermöglichen, wird zusätzlich ein Bindungspartner mit einem weiteren NMR-aktiven Kern wie z. B.  $^{15}\text{N}$  oder  $^{13}\text{C}$  markiert. Dies ermöglicht den Einsatz von mehrdimensionalen heteronuklearen Korrelations-NMR-Experimenten, die eine deutlich verbesserte Auflösung der individuellen Resonanzen zeigen. Mit Hilfe dieser Technik wurde die Bindung von  $\lambda$  N an AR1 (Einzelarbeit A und B) und  $\alpha$ CTD an AR2 untersucht (Einzelarbeit C). Für die Auswertung des CSP wird ein gewichteter Mittelwert aus den Differenzen der  $^1\text{H}$  und  $^{15}\text{N}$  chemischen Verschiebung von freiem und gebundenem Zustand verwendet. Dabei werden Änderungen, die einen Schwellenwert von  $\Delta\delta$  ( $^1\text{H}$ ,  $^{15}\text{N}$ ) von 0,04 ppm übersteigen, als signifikant angegeben und die entspre-

chenden Aminosäuren als an der Bindung teilnehmende Aminosäurereste klassifiziert (Bernd Meyer, 2003).

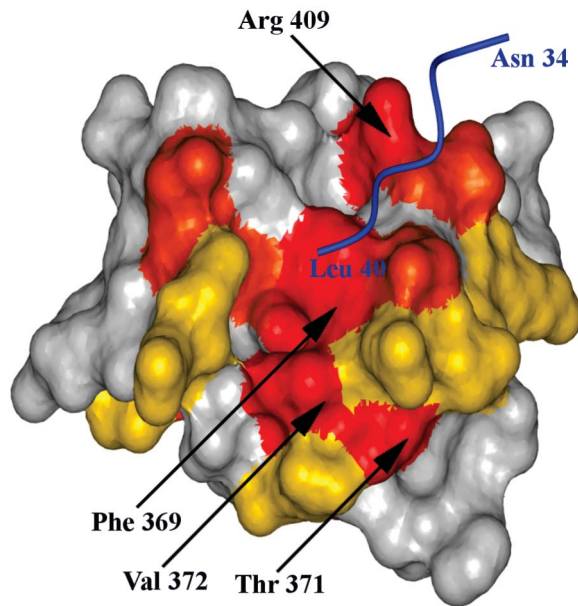
Darüber hinaus können mit Hilfe der NMR-Spektroskopie auch die Dissoziationskonstante  $K_D$  sowie die Stöchiometrie bestimmt werden. Aufgrund der für NMR-Experimente notwendigen Konzentrationen im  $\mu\text{M}$  bis  $\text{mM}$  Bereich können nur  $K_D$ -Werte in dieser Region quantitativ bestimmt werden. Besser eignet sich die NMR-Spektroskopie für eine stöchiometrische Titration, da hier die vorgelegte Konzentration über dem  $K_D$ -Wert liegen soll. So wurde sowohl für den  $\lambda$  N/AR1-Komplex (Einzelarbeit A) als auch für den AR2/ $\alpha$ CTD-Komplex (Einzelarbeit C) eine äquimolare Bindungskonstante bestimmt.

### 3.2. $\lambda$ N in Komplex mit der carboxyterminalen Domäne AR1 aus NusA

Da bei NMR-spektroskopischen Untersuchungen die Größe des Proteins oder Komplexes ein erhebliches Problem darstellt (Abschnitt 3.1), wurde beim  $\lambda$  N/NusA-Komplex versucht, die beteiligten Bindungsregionen soweit wie möglich einzugrenzen. Aus Affinitätschromatografiestudien und Röntgenstrukturanalysen war bekannt, dass nur der aminoterminal Bereich von  $\lambda$  N (Aminosäurereste 1 bis 53) sowie die AR1 Domäne von NusA für die Bindung eine Rolle spielen. Jedoch konnte nicht eindeutig bewiesen werden, ob AR2 an  $\lambda$  N bindet (Bonin *et al.*, 2004; Mogridge *et al.*, 1998).

Daher wurde zunächst die Bindung von  $\lambda$  N(1-53) an AR1 bzw. AR2 analysiert (Einzelarbeit A). CSP Analysen zeigten nur im Falle von AR1 signifikante Änderungen (Einzelarbeit A; Abb. 6). Aus der Titration ergab sich eine äquimolare Stöchiometrie für den  $\lambda$  N/AR1-Komplex.

Da in der Röntgenkristallstruktur nur die Aminosäurereste 34 bis 41 von  $\lambda$  N zu sehen sind (Abschnitt 1.4), sollte mit Hilfe der NMR-Spektroskopie geklärt werden, ob auch der Bereich 42-47 an AR1 bindet. CSP sowie NMR-Dynamikexperimente bewiesen einen Beitrag dieser Region an der Bindung. Eine Visualisierung der Veränderungen der chemischen Verschiebung auf der Oberfläche von AR1 zeigt, dass die Aminosäuren, welche die stärksten Veränderungen im CSP erfahren, direkt an der Bildung des Komplexes beteiligt sind (Abschnitt 1.4 und Abb. 3-1). Zusätzlich wird ein Bereich sichtbar, der von den Aminosäuren Phe-396, Val-372 und Thr-371 gebildet wird und eine Bindungsfläche für die Aminosäurereste 42-47 von  $\lambda$  N darstellen könnte.



**Abb. 3-1: Oberflächendarstellung des Komplexes von  $\lambda$  N(34-47) und AR1**

Die dargestellte Struktur basiert auf der gelösten Kristallstruktur (Bonin *et al.*, 2004). Das Peptidrückgrad von  $\lambda$  N(34-47) ist blau dargestellt. Die dem Lösungsmittel zugängliche Oberfläche von AR1 wurde mit verschiedenen Farben eingefärbt, basierend auf der Stärke des normalisierten CSP ( $\Delta_{\text{norm}}$ ):  $\Delta_{\text{norm}} < 0,09$  ppm in grau;  $0,09 \text{ ppm} < \Delta_{\text{norm}} < 0,15$  ppm in gelb;  $0,15 \text{ ppm} < \Delta_{\text{norm}} < 0,20$  ppm in orange; und  $0,20 \text{ ppm} < \Delta_{\text{norm}}$  in rot. Aminosäuren mit den stärksten Veränderungen in der chemischen Verschiebung sind durch Pfeile gekennzeichnet (Einzelarbeit A, Abb. 7).

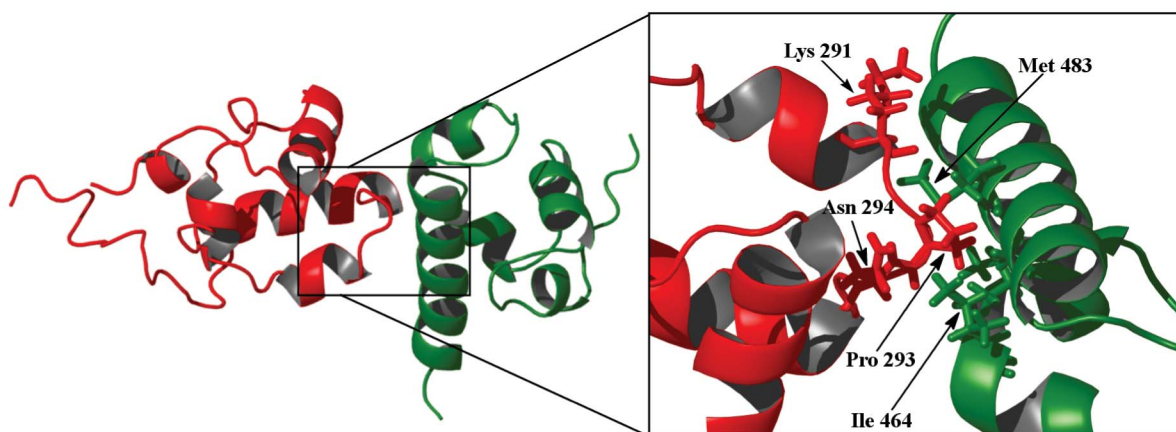
Überraschenderweise führt die Bindung von NusA-AR1 auch zu einer Konformationsänderung im Bereich des ARM von  $\lambda$  N (Abschnitt 1.3.1). Vor allem die Dynamikexperimente machen deutlich, dass durch die Bindung eine Helix induziert wird (Einzelarbeit A). Ob dies einen Einfluss auf die Bindung an die *nut boxB* RNA hat, wurde durch weitere NMR-Titrations geklärt. Der Vergleich der HSQC-Spektren des  $^{15}\text{N}$  markierten  $\lambda$  N(1-53) im Komplex mit unmarkiertem AR1 mit den Spektren des ternären Komplexes, bei dem zusätzlich die *nut boxB* RNA an AR1/ $\lambda$  N gebunden ist, zeigt allerdings keine Unterschiede. Somit treten beim Übergang vom binären in den ternären Komplex keine anderen strukturellen Änderungen im ARM von  $\lambda$  N auf (Einzelarbeit A). Dieser kooperative und modulare Aufbau von NusA-AR1/ $\lambda$  N, bei dem AR1 die Population an helikaler Konformation im ARM von  $\lambda$  N erhöht, erleichtert die nachfolgende Bindung weiterer Faktoren.

Analoges Verhalten wurde für die *nut boxB* RNA gezeigt (Johnson *et al.*, 2005). Bei der *nut boxB* RNA in Abwesenheit von  $\lambda$  N besteht ein Gleichgewicht zwischen freier und an  $\lambda$  N gebundener Konformation. Es wurde deswegen vorgeschlagen, dass die Erhöhung der Population von gebundener Konformation einen Mechanismus darstellt, um die Komplexbildung zu beschleunigen. In Analogie dazu könnte die Ausbildung der Helix durch die Bindung von AR1 an  $\lambda$  N auch die Erkennung der *nut boxB* RNA erleichtern. Die in dieser Arbeit neu gefundenen Erkenntnisse unterstützen damit das Modell, dass  $\lambda$  N zuerst an NusA und die RNAP bindet und erst danach mit der *nut* RNA in Wechselwirkung tritt (Xia *et al.*, 2003b).



### 3.3. NusA-AR2 bindet an die $\alpha$ CTD

In dem Komplex aus  $\lambda$  N/NusA-AR1 ist nur AR1 an der Bindung beteiligt (Abschnitt 3.2). Somit kann das ungebundene AR2 mit einem anderen Protein wechselwirken. Affinitätschromatografiestudien konnten eine Wechselwirkung zwischen NusA-AR2 und  $\alpha$ CTD nachweisen (Abschnitt 1.4) (Mah *et al.*, 2000; Einzelarbeit B). Im Rahmen dieser Arbeit wurde daher zuerst mit Hilfe von NMR-Titrationsen eine ausschließliche Bindung von AR2 in äquimolaren Mengen an  $\alpha$ CTD gezeigt (Einzelarbeit C). Im weiteren Verlauf dieser Arbeit wurde die Struktur des NusA-AR2/ $\alpha$ CTD-Komplexes bestimmt (Einzelarbeit C). Im Falle von AR2 bilden der Rest Ile-464 der Helix 4 und das Zentrum der Helix 5, bestehend aus Ile-482 und Met-483, eine hydrophobe Tasche. In diese bindet die Haarnadelschleife zwischen Helix 4 und Helix 5 von  $\alpha$ CTD (Aminosäurereste 291 bis 296) (Abb. 3-2). Das an der Spitze liegende Pro-293 taucht tief in die hydrophobe Tasche von AR2, wobei eine recht kleine Kontaktfläche von etwa  $800 \text{ \AA}^2$  entsteht. Daraus resultiert eine niedrige Affinität von  $\sim 10 \text{ }\mu\text{M}$ , die mittels Affinitätschromatografie und NMR-Titration geschätzt wurde (Mah *et al.*, 2000; Einzelarbeit C). Da NusA über die aminoternale Domäne an die RNAP gebunden ist, und somit sowohl AR2 als auch  $\alpha$ CTD in räumlicher Nähe sind, ist eine starke Wechselwirkung nicht erforderlich. Ein weiterer Grund für die schwache Affinität liegt in der biologischen Funktion. Um auf Signale schnell reagieren zu können, muss sich der Komplex NusA-AR2/ $\alpha$ CTD nicht nur bilden sondern auch lösen können, wozu eine schwache Affinität von Vorteil ist.



**Abb. 3-2: Struktur des Komplexes aus AR2 und  $\alpha$ CTD**

Die Abbildung zeigt die schematische Repräsentation der Struktur, mit der niedrigsten Energie. Die Haarnadelschleife von  $\alpha$ CTD (rot) interagiert mit den Helices 4 und 5 von AR2 (grün). Die vergrößerte Darstellung zeigt die für die Bindung wichtigsten Aminosäuren Lys-291, Pro-293 und Asn-294 auf Seiten der  $\alpha$ CTD, sowie die zu AR2 gehörenden Ile-464 und Met-483.

Die Bindung von AR2 an  $\alpha$ CTD bewirkt die Aufhebung der Autoinhibition von NusA (Abschnitt 1.4). Da das negativ geladene AR2 an die drei positiv geladenen RNA-Bindungsdomänen von NusA bindet, wird die Bindung von RNA-Sequenzen verhindert (Nudler und Gusarov, 2003). Erst durch die spezifische hydrophobe Wechselwirkung zwischen  $\alpha$ CTD und AR2 wird die RNA-Bindungsstelle freigegeben. Jetzt kann NusA mit der naszierenden RNA interagieren und so die Elongation regulieren, wie es in Abschnitt 1.4 beschrieben ist.

Auf Seiten der  $\alpha$ CTD wird durch AR2 die „265-Determinante“ blockiert, die für die Wechselwirkung von  $\alpha$ CTD mit dem UP-Element verantwortlich ist (Abschnitt 1.2.2). Diese Blockade wird ebenso beim SoxS System gefunden. SoxS bindet an die „265-Determinante“ von  $\alpha$ CTD und verhindert so die Wechselwirkung zwischen  $\alpha$ CTD und dem UP-Element (Abschnitt 1.2.2). Dies verhindert eine Interaktion der RNAP mit unspezifischen Promotoren. Die durch SoxS veränderte RNAP erkennt nun ausschließlich SoxS spezifische Promotoren, wodurch Gene aktiviert werden, die den oxidativen Stress in der Zelle bekämpfen. In Analogie dazu verdrängt AR2 durch Bindung an die „265-Determinante“  $\alpha$ CTD vom UP-Element. So wird die RNAP von der Initiationsphase in die Elongationsphase überführt. Da die beiden anderen regulatorischen Determinanten von  $\alpha$ CTD frei zugänglich sind, können sowohl Aktivatorproteine („287-Determinante“) als auch der  $\sigma^{70}$ -Faktor („261-Determinante“) weiterhin mit der RNAP interagieren, wie es beim stochastischen Modell vorgeschlagen wird (Abschnitt 1.2.2).

### 3.4. Vergleich von AR1 und AR2

Die beiden NusA-Domänen AR1 und AR2 weisen sowohl eine hohe Sequenz- als auch Strukturhomologie auf (Einzelarbeit B). Trotzdem interagieren beide mit unterschiedlichen Proteinen (Abschnitt 3.2 und 3.4). Im Komplex von AR1 mit  $\lambda$  N zeigt AR1 einen hydrophoben Bereich zwischen Helix 2 und dem Loop zwischen Helix 4 und 5. Vor allem Leu-40 von  $\lambda$  N bildet starke Kontakte mit diesem hydrophoben Bereich in AR1 aus (Abschnitt 1.4 und Abb. 1-8). Ebenso wird die Interaktion bei AR2/ $\alpha$ CTD durch hydrophobe Kontakte im Bereich der Helices 4 und 5 von AR2 vermittelt. In diesem Fall wird aber der hydrophobe Bereich zwischen dem aminoterminalen Bereich der Helix 4 und dem Zentrum der Helix 5 gefunden. Der Bereich in AR2 befindet sich im Wesentlichen auf der gegenüberliegenden Seite im Vergleich zu AR1/ $\lambda$  N (Einzelarbeit C). In beiden Fällen sind die relevanten Sequenzpositionen für die jeweilige Interaktion nicht konserviert. Die Spezifität von  $\lambda$  N zu AR1 basiert auf dem hydro-

phoben Val-372 von AR1. In AR2 wird an der korrespondierenden Stelle ein geladenes Lys-447 gefunden, das die für den AR1/ $\lambda$  N-Komplex notwendigen hydrophoben Kontakte verhindern würde. Wird die Bindungsstelle von AR2 für  $\alpha$ CTD auf AR1 projiziert, wird die Anwesenheit von Arg-407 und Glu-408 in AR1 anstelle von Ile-482 und Met-483 in AR2 ersichtlich. Die geladenen Aminosäuren in AR1 würden die notwendigen hydrophoben Wechselwirkungen zerstören. Deswegen kann keine Bindung zwischen AR1 und  $\alpha$ CTD stattfinden (Einzelarbeit C).

### 3.5. Wechselwirkung zwischen NusA und RNA

Aufgrund der Autoinhibition durch AR2 kann das freie NusA keine RNA binden (Abschnitt 1.4). Deshalb wurde ein NusA Konstrukt ( $\Delta$ NusA) verwendet, bei dem die carboxy-terminalen Domänen AR1 und AR2 fehlen. Indem AR2 entweder durch die Bindung an RNAP oder durch Deletion entfernt wird, wird eine Interaktion mit der RNA ermöglicht (Arnvig *et al.*, 2004; Mah *et al.*, 2000). *E. coli* Stämme mit einem um AR1/AR2 verkürztem NusA sind überlebensfähig (Bylund *et al.*, 2001).

Mit Fluoreszenzanisotropie-Messungen wurde die Bindung an die Antiterminationssignale im *rrn* Operon und an die Signalsequenzen *nutL* und *nutR* des Phagen  $\lambda$  untersucht (Einzelarbeit D). Die drei Signale zeigen *in vivo* unterschiedliche Eigenschaften. So wird im Falle des *rrn* Operon kein zusätzlicher Proteinfaktor benötigt, um Antitermination hervorzurufen, dafür wird nur an Rho-abhängigen Signalen die Transkription fortgesetzt (Abschnitt 1.3.2). Im Gegensatz dazu wird Transkription im Genom des Phagen  $\lambda$  sowohl an Rho-abhängigen als auch an intrinsischen Terminatoren fortgeführt, hierfür wird  $\lambda$  N benötigt (Abschnitt 1.3.1). *In vivo* unterscheiden sich *nutL* und *nutR* nur durch unterschiedliche Effizienzen in der Antitermination (persönliche Mitteilung Gottesman).

Die von mir durchgeführten Fluoreszenzstudie zeigt bei allen drei untersuchten Terminationsignalen eine Bindung von  $\Delta$ NusA ausschließlich an den *Spacer*, der *boxB* von *boxA* trennt (Einzelarbeit D). Dieses Ergebnis wurde durch NMR-Titrations und Mutationen in der *boxA* und im *Spacer* validiert. So zeigen nur Mutationen im Bereich des *Spacers* eine signifikante Änderung des  $K_D$ -Wertes, während sich Mutationen im Bereich der *boxA* nicht auf die Bindung auswirkten. Dies korrespondiert mit *in vitro* und *in vivo* Daten, bei denen Mutationen oder Deletionen der *boxA* die Antiterminationsrate unverändert ließen (Patterson *et al.*, 1994; Zuber *et al.*, 1987). Da  $\Delta$ NusA an den *Spacer* bindet, besteht somit keine Konkurrenz um die Bindung an die *boxA* zwischen NusA und dem Komplex aus NusB und NusE, bei denen eine

Bindung an die *boxA* nachgewiesen wurde (Greive *et al.*, 2005). In der Fluoreszenzstudie wurden unterschiedliche  $K_D$ -Werte für *rrn*, *nutL* und *nutR* RNA gefunden. Für *rrn* wird die höchste und für *nutR* die niedrigste Affinität gemessen (Einzelarbeit D). Dies erklärt, warum im Falle der *rrn* Antitermination keine zusätzlichen Proteine wie  $\lambda$  N notwendig sind, welche die Stabilität des Antiterminationskomplexes erhöhen. Die verschiedenen  $K_D$ -Werte zwischen *nutL* und *nutR* bestätigen so unterschiedliche Antiterminationseffizienzen *in vivo* und *in vitro* (persönliche Mitteilung Gottesman). Generell zeigen alle drei gemessenen RNAs nur  $K_D$ -Werte im  $\mu$ M Bereich. Die räumliche Nähe von NusA, das an die RNAP gebunden ist, zur RNA könnte die schwache Bindung erklären.

### 3.6. Aufbau des Antiterminationskomplexes

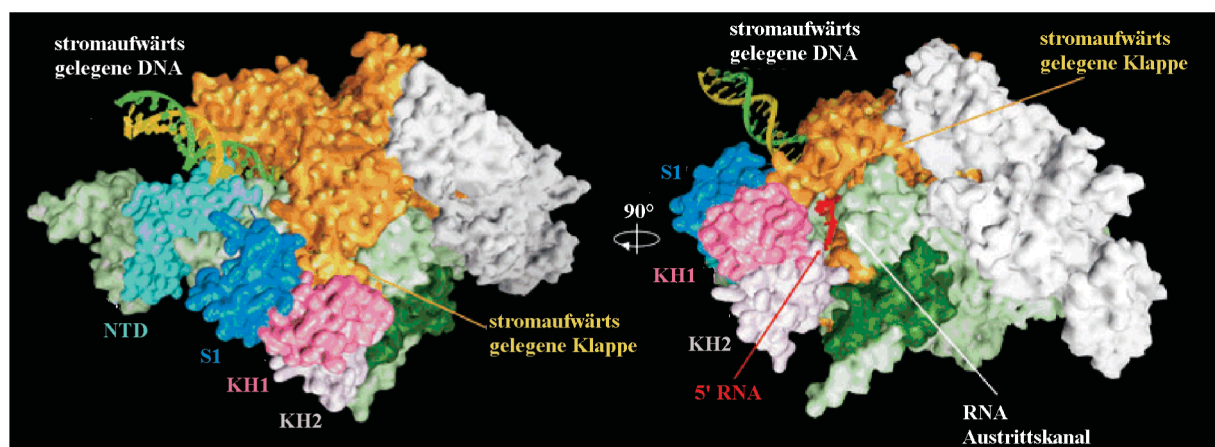
Verschiedene mechanistische Modelle versuchen, die Funktion von NusA anhand biochemischer und genetischer Studien zu erklären (Nudler und Gottesman, 2002). NusA konkurriert mit dem  $\sigma^{70}$ -Faktor. So besitzt der  $\sigma^{70}$ -Faktor eine höhere Affinität zum Initiationskomplex ( $K_D \approx 10^{-10}$  M) als zum Elongationskomplex ( $K_D \approx 10^{-6}$  M), während NusA mit einem höherem  $K_D$ -Wert ( $K_D \approx 10^{-8}$  M) als der  $\sigma^{70}$ -Faktor an den Elongationskomplex bindet (Greenblatt und Li, 1981; Mooney *et al.*, 2005). Die Konkurrenz wird durch *footprint* Analysen bestätigt, bei denen NusA an die gleiche Stelle wie der  $\sigma^{70}$ -Faktor bindet (Traviglia *et al.*, 1999). Die in dieser Arbeit gefundenen Affinitäten von  $\Delta$ NusA an RNA und an  $\alpha$ CTD sind niedriger und lassen auf zusätzliche Interaktion von NusA mit der RNAP hindeuten, an denen die NTD von NusA beteiligt sein könnte.

Da der NTD starke strukturelle Homologien mit der Region 2 aus  $\sigma^{70}$  aufweist, kann die NusA Struktur, die für *T. maritima* und *M. tuberculosis* bestimmt wurde, mit der bekannten Struktur der RNAP im Komplex mit dem  $\sigma^{70}$ -Faktor überlagert werden (Abb. 3-3), indem der NTD von NusA mit der Region 2 von  $\sigma^{70}$  abgeglichen wird (Borukhov *et al.*, 2005). Die genaue Lage der S1-, KH1 und KH2-Domänen hängt jedoch von der jeweils verwendeten Struktur ab. Wird die Struktur aus *T. maritima* zur Modellierung verwendet, liegen die drei Domänen oberhalb der Klappendomäne, im Falle der *M. tuberculosis* Struktur unterhalb (Borukhov *et al.*, 2005). Da letztere besser die Vernetzungsexperimente beschreiben (Gusarov und Nudler, 2001; Touloukhonov *et al.*, 2001), wurde diese in Abb. 3-3 verwendet.

In dem Modell würde die positiv geladene NTD von NusA den nicht codierenden Strang der stromaufwärts gelegenen DNA oder Teile des RNA-DNA Hybrides binden (Abb. 3-3), umso das Pausieren der Transkription zu steuern (Borukhov *et al.*, 2005). Die darauf folgenden S1

und KH Domänen, die in der Nähe der Klappendomäne lokalisiert sind, erlauben eine weitere Interaktion mit  $\lambda$  N, NusB, NusE oder NusG. Ferner deutet das Homologiemodell auf eine Wechselwirkung zwischen den beiden KH Domänen und dem 5' Ende der RNA hin (Abb. 3-3) und steht so im Einklang mit der Struktur von NusA aus *M. tuberculosis* gebunden an die *rrn boxC*, bei der die beiden KH-Domänen spezifische Kontakte mit der RNA ausbilden (Arnvig *et al.*, 2004; Beuth *et al.*, 2005).

Die räumliche Nähe von NusA zur RNA stimmt ebenso mit dem in Abschnitt 1.4 vorgestellten Modell überein, bei dem NusA durch spezifische Wechselwirkung mit dem 5' Teil der RNA zum einen die Termination zum anderen die Antitermination fördert. Da in den kristallografisch bestimmten NusA Strukturen die beiden carboxyterminalen Domänen AR1 und AR2 fehlen, können diese im Modell nicht dargestellt werden. Die in dieser Arbeit gewonnen Erkenntnisse können trotzdem anhand des Modells erklärt werden. Im Anschluss an die NusA-KH2-Domäne folgt AR1, das über einen flexiblen Linker verbunden ist. Somit wäre AR1 in räumlicher Nähe zum 5' Ende der RNA. Indem  $\lambda$  N zuerst an AR1 bindet, wird es in die Nähe des RNA-Austrittskanals gebracht. Dort kann  $\lambda$  N das 5' Ende der RNA nach der *boxB* Sequenz abtasten, und diese nach erfolgter Transkription sofort binden. Zusätzlich würde der  $\lambda$  N/AR1 Komplex das intrinsisch unstrukturierte  $\lambda$  N Protein vor einer Degradierung schützen (Abschnitt 1.3.1). Die nachfolgende AR2 Domäne könnte ferner an die  $\alpha$ CTD gebunden sein, da sowohl AR2 als auch  $\alpha$ CTD über längere flexible Linker verbunden sind.



**Abb. 3-3: Strukturmodell von NusA im Komplex mit der RNAP**

Das Strukturmodell basiert auf der Kristallstruktur der RNAP aus *T. thermophilus* (Vassylyev *et al.*, 2002) und wurde von Borukhov *et al.* (2005) übernommen. Die NTD (cyan) bindet in der Nähe der stromaufwärtsgelegenen DNA (grün und gelb), während die drei RNA-Bindungsdomänen S1 (blau), KH1 (rosa) und KH2 (hellrosa) an die  $\beta$  (orange) und  $\beta'$  (grün) binden. Die NusA Struktur wurde aus *T. maritima* sowie *M. tuberculosis* übernommen (Gopal *et al.*, 2001; Borukhov *et al.*, 2005; Worbs *et al.*, 2001). Das 5' Ende der RNA (rot) liegt im Modell in räumlicher Nähe zu den beiden KH Domänen.

#### 4. Abkürzungsverzeichnis

AR	Acidic repeat
ATP	Adenosintriphosphat
CAP	Catabolite activator protein
CSP	Chemical shift perturbation
DNA	Desoxyribonukleinsäure
dwDNA	Downstream DNA
<i>E. coli</i>	<i>Escherichia coli</i>
GTP	Guanosintriphosphat
HSQC	Heteronuclear single quantum coherence
KH	K homologe Domäne
<i>M. tuberculosis</i>	<i>Mycobakterium tuberculosis</i>
NMR	Nuclear magnetic resonance spectroscopy
nt	Nukleotide
NTD	Aminoterminal Domäne von NusA
Nus A, B, E, G	N utilization substance A, B, E, G
<i>nut</i>	N utilization site
OxyR	Oxidative stress regulator protein
RNA	Ribonukleinsäure
RNAP	RNA-Polymerase
<i>rrn</i>	Ribosomale RNA
<i>rut</i>	Rho utilization
S1	S homologe Domäne 1
<i>T. maritima</i>	<i>Thermotoga maritima</i>
<i>T. thermophilus</i>	<i>Thermus thermophilus</i>
TEC	Transkriptionskomplex
TH	Trigger Helix
TL	Trigger Loop
UP-element	Upstream promotor element
$\alpha$ CTD	Carboxytermincarboxyterminale Domäne der $\alpha$ -Untereinheit der RNAP
$\Delta$ NusA	Aminosäurereste 132 bis 348 von NusA

## 5. Literaturverzeichnis

- Albrechtsen B, Squires CL, Li S und Squires C** (1990) Antitermination of characterized transcriptional terminators by the Escherichia coli rrnG leader region. *J Mol Biol* **213**: 123-134
- Altieri AS, Mazzulla MJ, Horita DA, Coats RH, Wingfield PT, Das A, Court DL und Byrd RA** (2000) The structure of the transcriptional antiterminator NusB from Escherichia coli. *Nat Struct Biol* **7**: 470-474
- Arnvig KB, Pennell S, Gopal B und Colston MJ** (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
- Banerjee S, Chalissery J, Bandey I und Sen R** (2006) Rho-dependent transcription termination: more questions than answers. *J Microbiol* **44**: 11-22
- Benoff B, Yang H, Lawson CL, Parkinson G, Liu J, Blatter E, Ebright YW, Berman HM und Ebright RH** (2002) Structural basis of transcription activation: the CAP-alpha CTD-DNA complex. *Science* **297**: 1562-1566
- Bernd Meyer TP**, (2003) NMR-Techniken zum Screening und zur Identifizierung der Bindung von Liganden an Proteinrezeptoren. *Angewandte Chemie* **115**: 890-918
- Beuth B, Pennell S, Arnvig KB, Martin SR und Taylor IA** (2005) Structure of a Mycobacterium tuberculosis NusA-RNA complex. *EMBO J* **24**: 3576-3587
- Bonin I** (2004) Structural and Biochemical Studies of Two Nus Family Proteins: NusB and NusA AR1-N Complex. Doktorarbeit TU München
- Bonin I, Muhlberger R, Bourenkov GP, Huber R, Bacher A, Richter G und Wahl MC** (2004) Structural basis for the interaction of Escherichia coli NusA with protein N of phage lambda. *Proc Natl Acad Sci U S A* **101**: 13762-13767
- Borukhov S, Lee J und Laptenko O** (2005) Bacterial transcription elongation factors: new insights into molecular mechanism of action. *Mol Microbiol* **55**: 1315-1324
- Borukhov S, Severinov K** (2002) Role of the RNA polymerase sigma subunit in transcription initiation. *Res Microbiol* **153**: 557-562
- Buck M, Gallegos MT, Studholme DJ, Guo Y und Gralla JD** (2000) The bacterial enhancer-dependent sigma(54) (sigma(N)) transcription factor. *J Bacteriol* **182**: 4129-4136
- Bylund GO, Lovgren JM und Wikstrom PM** (2001) Characterization of mutations in the metY-nusA-infB operon that suppress the slow growth of a DeltarimM mutant. *J Bacteriol* **183**: 6095-6106
- Ciampi MS** (2006) Rho-dependent terminators and transcription termination. *Microbiology* **152**: 2515-2528
- Cilley CD, Williamson JR** (1997) Analysis of bacteriophage N protein and peptide binding to boxB RNA using polyacrylamide gel coelectrophoresis (PACE). *RNA* **3**: 57-67

- Clarkson J, Campbell ID** (2003) Studies of protein-ligand interactions by NMR. *Biochem Soc Trans* **31**: 1006-1009
- Clore GM, Gronenborn AM** (1998) Determining the structures of large proteins and protein complexes by NMR. *Trends Biotechnol* **16**: 22-34
- Condon C, French S, Squires C und Squires CL** (1993) Depletion of functional ribosomal RNA operons in *Escherichia coli* causes increased expression of the remaining intact copies. *EMBO J* **12**: 4305-4315
- Condon C, Philips J, Fu ZY, Squires C und Squires CL** (1992) Comparison of the expression of the seven ribosomal RNA operons in *Escherichia coli*. *EMBO J* **11**: 4175-4185
- Condon C, Squires C und Squires CL** (1995) Control of rRNA transcription in *Escherichia coli*. *Microbiol Rev* **59**: 623-645
- Court DL, Oppenheim AB und Adhya SL** (2007) A new look at bacteriophage lambda genetic networks. *J Bacteriol* **189**: 298-304
- Cramer P, Bushnell DA und Kornberg RD** (2001) Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. *Science* **292**: 1863-1876
- Dangi B, Gronenborn AM, Rosner JL und Martin RG** (2004) Versatility of the carboxy-terminal domain of the alpha subunit of RNA polymerase in transcriptional activation: use of the DNA contact site as a protein contact site for MarA. *Mol Microbiol* **54**: 45-59
- Darst S** (2001) Bacterial RNA polymerase. *Current Opinion in Structural Biology*, **11**: 155-162
- Das A, Pal M, Mena JG, Whalen W, Wolska K, Crossley R, Rees W, von Hippel PH, Costantino N, Court D, Mazzulla M, Altieri AS, Byrd RA, Chattopadhyay S, DeVito J und Ghosh B** (1996) Components of multiprotein-RNA complex that controls transcription elongation in *Escherichia coli* phage lambda. *Methods Enzymol* **274**: 374-402
- Dodd IB, Shearwin KE und Egan JB** (2005) Revisited gene regulation in bacteriophage lambda. *Curr Opin Genet Dev* **15**: 145-152
- Drahos D, Galluppi GR, Caruthers M und Szybalski W** (1982) Synthesis of the nutL DNA segments and analysis of antitermination and termination functions in coliphage lambda. *Gene* **18**: 343-354
- Eisenmann A, Schwarz S, Prasch S, Schweimer K und Rosch P** (2005) The *E. coli* NusA carboxy-terminal domains are structurally similar and show specific RNAP- and lambdaN interaction. *Protein Sci* **14**: 2018-2029
- Estrem ST, Gaal T, Ross W und Gourse RL** (1998) Identification of an UP element consensus sequence for bacterial promoters. *Proc Natl Acad Sci U S A* **95**: 9761-9766
- Estrem ST, Ross W, Gaal T, Chen ZW, Niu W, Ebright RH und Gourse RL** (1999) Bacterial promoter architecture: subsite structure of UP elements and interactions with the carboxy-terminal domain of the RNA polymerase alpha subunit. *Genes Dev* **13**: 2134-2147



- Fish RN, Kane CM** (2002) Promoting elongation with transcript cleavage stimulatory factors. *Biochim Biophys Acta* **1577**: 287-307
- Franklin NC** (1993) Clustered Arginine Residues of Bacteriophage  $\lambda$  N Protein are Essential to Antitermination of Transcription, but Their Locale Cannot Compensate for boxB Loop Defects. *Journal of Molecular Biology*, **231**: 343-360
- Friedman DI, Baron LS** (1974) Genetic characterization of a bacterial locus involved in the activity of the N function of phage lambda. *Virology* **58**: 141-148
- Friedman DI, Court DL** (2001) Bacteriophage lambda: alive and well and still doing its thing. *Curr Opin Microbiol* **4**: 201-207
- Friedman DI, Schauer AT, Baumann MR, Baron LS und Adhya SL** (1981) Evidence that ribosomal protein S10 participates in control of transcription termination. *Proc Natl Acad Sci U S A* **78**: 1115-1118
- Gnatt AL, Cramer P, Fu J, Bushnell DA und Kornberg RD** (2001) Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 Å resolution. *Science* **292**: 1876-1882
- Gopal B, Haire LF, Gamblin SJ, Dodson EJ, Lane AN, Papavinasasundaram KG, Colston MJ und 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
- Gotta SL, Miller OL, Jr und French SL** (1991) rRNA transcription rate in *Escherichia coli*. *J Bacteriol* **173**: 6647-6649
- Gottesman ME, Weisberg RA** (2004) Little lambda, who made thee? *Microbiol Mol Biol Rev* **68**: 796-813
- Gourse RL, de Boer HA und Nomura M** (1986) DNA determinants of rRNA synthesis in *E. coli*: growth rate dependent regulation, feedback inhibition, upstream activation, antitermination. *Cell* **44**: 197-205
- Greenblatt J, Li J** (1981) 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 TF, Legault P, Mogridge J, Li J und Kay LE** (1998) Structure and mechanism in transcriptional antitermination by the bacteriophage lambda N protein. *Cold Spring Harb Symp Quant Biol* **63**: 327-336
- Greenblatt J, Malnoe P und Li J** (1980) Purification of the gene N transcription anti-termination protein of bacteriophage lambda. *J Biol Chem* **255**: 1465-1470
- Greive SJ, Lins AF und von Hippel PH** (2005) Assembly of an RNA-protein complex. Binding of NusB and NusE (S10) proteins to boxA RNA nucleates the formation of the anti-termination complex involved in controlling rRNA transcription in *Escherichia coli*. *J Biol Chem* **280**: 36397-36408
- Greive SJ, von Hippel PH** (2005) Thinking quantitatively about transcriptional regulation. *Nat Rev Mol Cell Biol* **6**: 221-232

- Gusarov I, Nudler E** (2001) Control of intrinsic transcription termination by N and NusA: the basic mechanisms. *Cell* **107**: 437-449
- Gusarov I, Nudler E** (1999) The mechanism of intrinsic transcription termination. *Mol Cell* **3**: 495-504
- Hasan N, Szybalski W** (1986) Effect of the promoter structure on the nutL transcription anti-termination function. *Gene*, **50**: 97-100
- Heinrich T, Condon C, Pfeiffer T und Hartmann RK** (1995) Point mutations in the leader boxA of a plasmid-encoded *Escherichia coli* rrnB operon cause defective antitermination in vivo. *J Bacteriol* **177**: 3793-3800
- Helmann JD** (1999) Anti-sigma factors. *Curr Opin Microbiol* **2**: 135-141
- Henkin TM** (2000) Transcription termination control in bacteria. *Curr Opin Microbiol* **3**: 149-153
- Henkin TM** (1996) Control of transcription termination in prokaryotes. *Annu Rev Genet* **30**: 35-57
- Horwitz RJ, Li J und Greenblatt J** (1987) An elongation control particle containing the N gene transcriptional antitermination protein of bacteriophage lambda. *Cell* **51**: 631-641
- Hughes KT, Mathee K** (1998) The anti-sigma factors. *Annu Rev Microbiol* **52**: 231-286
- Jeon YH, Yamazaki T, Otomo T, Ishihama A und Kyogoku Y** (1997) Flexible linker in the RNA polymerase alpha subunit facilitates the independent motion of the C-terminal activator contact domain. *J Mol Biol* **267**: 953-962
- Johnson NP, Baase WA und von Hippel PH** (2005) Low energy CD of RNA hairpin unveils a loop conformation required for lambdaN antitermination activity. *J Biol Chem* **280**: 32177-32183
- Jucker FM, Heus HA, Yip PF, Moors EHM und Pardi A** (1996) A Network of Heterogeneous Hydrogen Bonds in GNRA Tetraloops. *Journal of Molecular Biology*, **264**: 968-980
- Komissarova N, Becker J, Solter S, Kireeva M und Kashlev M** (2002) Shortening of RNA:DNA hybrid in the elongation complex of RNA polymerase is a prerequisite for transcription termination. *Mol Cell* **10**: 1151-1162
- Korzheva N, Mustaev A, Kozlov M, Malhotra A, Nikiforov V, Goldfarb A und Darst SA** (2000) A structural model of transcription elongation. *Science* **289**: 619-625
- Landick R** (2006) The regulatory roles and mechanism of transcriptional pausing. *Biochem Soc Trans* **34**: 1062-1066
- Legault P, Li J, Mogridge J, Kay LE und Greenblatt J** (1998) NMR structure of the bacteriophage lambda N peptide/boxB RNA complex: recognition of a GNRA fold by an arginine-rich motif. *Cell* **93**: 289-299
- Li J, Mason SW und Greenblatt J** (1993) Elongation factor NusG interacts with termination factor rho to regulate termination and antitermination of transcription. *Genes Dev* **7**: 161-172

- Li SC, Squires CL und Squires C** (1984) Antitermination of *E. coli* rRNA transcription is caused by a control region segment containing lambda nut-like sequences. *Cell* **38**: 851-860
- Liu K, Hanna MM** (1995) NusA contacts nascent RNA in *Escherichia coli* transcription complexes. *J Mol Biol* **247**: 547-558
- Liu K, Zhang Y, Severinov K, Das A und Hanna MM** (1996) Role of *Escherichia coli* RNA polymerase alpha subunit in modulation of pausing, termination and anti-termination by the transcription elongation factor NusA. *EMBO J* **15**: 150-161
- Mah TF, Kuznedelov K, Mushegian A, Severinov K und Greenblatt J** (2000) The alpha subunit of *E. coli* RNA polymerase activates RNA binding by NusA. *Genes Dev* **14**: 2664-2675
- Mah TF, Li J, Davidson AR und Greenblatt J** (1999) Functional importance of regions in *Escherichia coli* elongation factor NusA that interact with RNA polymerase, the bacteriophage lambda N protein and RNA. *Mol Microbiol* **34**: 523-537
- Mason SW, Greenblatt J** (1991) Assembly of transcription elongation complexes containing the N protein of phage lambda and the *Escherichia coli* elongation factors NusA, NusB, NusG, and S10. *Genes Dev* **5**: 1504-1512
- Mason SW, Li J und Greenblatt J** (1992) Host factor requirements for processive antitermination of transcription and suppression of pausing by the N protein of bacteriophage lambda. *J Biol Chem* **267**: 19418-19426
- Maurizi M** (1987) Degradation in vitro of bacteriophage lambda N protein by Lon protease from *Escherichia coli*. *J Biol Chem* **262**: 2696-2703
- Metzger W, Schickor P, Meier T, Werel W und Heumann H** (1993) Nucleation of RNA Chain Formation by *Escherichia coli* DNA-dependent RNA Polymerase. *Journal of Molecular Biology*, **232**: 35-49
- Mogridge J, Legault P, Li J, Van Oene MD, Kay LE und Greenblatt J** (1998) Independent ligand-induced folding of the RNA-binding domain and two functionally distinct antitermination regions in the phage lambda N protein. *Mol Cell* **1**: 265-275
- Mooney RA, Darst SA und Landick R** (2005) Sigma and RNA polymerase: an on-again, off-again relationship? *Mol Cell* **20**: 335-345
- Morgan EA** (1986) Antitermination mechanisms in rRNA operons of *Escherichia coli*. *J Bacteriol* **168**: 1-5
- Mukherjee K, Chatterji D** (1997) Studies on the omega subunit of *Escherichia coli* RNA polymerase--its role in the recovery of denatured enzyme activity. *Eur J Biochem* **247**: 884-889
- Mukherjee K, Nagai H, Shimamoto N und Chatterji D** (1999) GroEL is involved in activation of *Escherichia coli* RNA polymerase devoid of the omega subunit in vivo. *Eur J Biochem* **266**: 228-235

- Murakami K, Fujita N und Ishihama A** (1996) Transcription factor recognition surface on the RNA polymerase alpha subunit is involved in contact with the DNA enhancer element. *EMBO J* **15**: 4358-4367
- Murakami KS, Darst SA** (2003) Bacterial RNA polymerases: the whole story. *Curr Opin Struct Biol* **13**: 31-39
- Neuman KC, Abbondanzieri EA, Landick R, Gelles J und Block SM** (2003) Ubiquitous transcriptional pausing is independent of RNA polymerase backtracking. *Cell* **115**: 437-447
- Nudler E** (1999) Transcription elongation: structural basis and mechanisms. *J Mol Biol* **288**: 1-12
- Nudler E, Gottesman ME** (2002) Transcription termination and anti-termination in *E. coli*. *Genes Cells* **7**: 755-768
- Nudler E, Gusarov I** (2003) Analysis of the intrinsic transcription termination mechanism and its control. *Methods Enzymol* **371**: 369-382
- Paget MS, Helmann JD** (2003) The sigma70 family of sigma factors. *Genome Biol* **4**: 203
- Patterson TA, Zhang Z, Baker T, Johnson LL, Friedman DI und Court DL** (1994) Bacteriophage lambda N-dependent transcription antitermination. Competition for an RNA site may regulate antitermination. *J Mol Biol* **236**: 217-228
- Ptashne M** (2006) Lambda's switch: lessons from a module swap. *Curr Biol* **16**: R459-62
- Rees WA, Weitzel SE, Yager TD, Das A und von Hippel PH** (1996) Bacteriophage lambda N protein alone can induce transcription antitermination in vitro. *Proc Natl Acad Sci U S A* **93**: 342-346
- Rosenberg M, Court D, Shimatake H, Brady C und Wulff DL** (1978) The relationship between function and DNA sequence in an intercistronic regulatory region in phage lambda. *Nature* **272**: 414-423
- Ross W, Gosink KK, Salomon J, Igarashi K, Zou C, Ishihama A, Severinov K und Gourse RL** (1993) A third recognition element in bacterial promoters: DNA binding by the alpha subunit of RNA polymerase. *Science* **262**: 1407-1413
- Savery NJ, Lloyd GS, Kainz M, Gaal T, Ross W, Ebright RH, Gourse RL und Busby SJ** (1998) Transcription activation at Class II CRP-dependent promoters: identification of determinants in the C-terminal domain of the RNA polymerase alpha subunit. *EMBO J* **17**: 3439-3447
- Scharpf M, Sticht H, Schweimer K, Boehm M, Hoffmann S und Rosch P** (2000) Antitermination in bacteriophage lambda. The structure of the N36 peptide-boxB RNA complex. *Eur J Biochem* **267**: 2397-2408
- Schneider D, Gold L und Platt T** (1993) Selective enrichment of RNA species for tight binding to *Escherichia coli* rho factor. *FASEB J* **7**: 201-207
- Severinov K** (2000) RNA polymerase structure-function: insights into points of transcriptional regulation. *Curr Opin Microbiol* **3**: 118-125

- Shah IM, Wolf RE, Jr** (2004) Novel protein--protein interaction between Escherichia coli SoxS and the DNA binding determinant of the RNA polymerase alpha subunit: SoxS functions as a co-sigma factor and redeploys RNA polymerase from UP-element-containing promoters to SoxS-dependent promoters during oxidative stress. *J Mol Biol* **343**: 513-532
- Shilatifard A, Conaway RC und Conaway JW** (2003) The RNA polymerase II elongation complex. *Annu Rev Biochem* **72**: 693-715
- Shimamoto N, Kamigochi T und Utiyama H** (1986) Release of the sigma subunit of Escherichia coli DNA-dependent RNA polymerase depends mainly on time elapsed after the start of initiation, not on length of product RNA. *J Biol Chem* **261**: 11859-11865
- Skordalakes E, Berger JM** (2003) Structure of the Rho transcription terminator: mechanism of mRNA recognition and helicase loading. *Cell* **114**: 135-146
- Squires CL, Greenblatt J, Li J, Condon C und Squires CL** (1993) Ribosomal RNA anti-termination in vitro: requirement for Nus factors and one or more unidentified cellular components. *Proc Natl Acad Sci U S A* **90**: 970-974
- Su L, Radek JT, Hallenga K, Hermanto P, Chan G, Labeots LA und Weiss MA** (1997) RNA recognition by a bent alpha-helix regulates transcriptional antitermination in phage lambda. *Biochemistry* **36**: 12722-12732
- Tan R, Frankel AD** (1995) Structural variety of arginine-rich RNA-binding peptides. *Proc Natl Acad Sci U S A* **92**: 5282-5286
- Tao K, Fujita N und Ishihama A** (1993) Involvement of the RNA polymerase alpha subunit C-terminal region in co-operative interaction and transcriptional activation with OxyR protein. *Mol Microbiol* **7**: 859-864
- Touloukhonov I, Artsimovitch I und Landick R** (2001) Allosteric control of RNA polymerase by a site that contacts nascent RNA hairpins. *Science* **292**: 730-733
- Touloukhonov I, Landick R** (2003) The flap domain is required for pause RNA hairpin inhibition of catalysis by RNA polymerase and can modulate intrinsic termination. *Mol Cell* **12**: 1125-1136
- Travaglia SL, Datwyler SA, Yan D, Ishihama A und Meares CF** (1999) Targeted protein footprinting: where different transcription factors bind to RNA polymerase. *Biochemistry* **38**: 15774-15778
- Uptain SM, Kane CM und Chamberlin MJ** (1997) Basic mechanisms of transcript elongation and its regulation. *Annu Rev Biochem* **66**: 117-172
- Van Gilst MR, von Hippel PH** (1997) Assembly of the N-dependent antitermination complex of phage lambda: NusA and RNA bind independently to different unfolded domains of the N protein. *J Mol Biol* **274**: 160-173
- Vassilyev DG, Sekine S, Laptenko O, Lee J, Vassilyeva MN, Borukhov S und Yokoyama S** (2002) Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 Å resolution. *Nature* **417**: 712-719

- Vassilyev DG, Vassilyeva MN, Perederina A, Tahirov TH und Artsimovitch I** (2007a) Structural basis for transcription elongation by bacterial RNA polymerase. *Nature* **448**: 157-162
- Vassilyev DG, Vassilyeva MN, Zhang J, Palangat M, Artsimovitch I und Landick R** (2007b) Structural basis for substrate loading in bacterial RNA polymerase. *Nature* **448**: 163-168
- Vogel U, Jensen KF** (1995) Effects of the antiterminator BoxA on transcription elongation kinetics and ppGpp inhibition of transcription elongation in *Escherichia coli*. *J Biol Chem* **270**: 18335-18340
- Vogel U, Jensen KF** (1994) Effects of guanosine 3',5'-bisdiphosphate (ppGpp) on rate of transcription elongation in isoleucine-starved *Escherichia coli*. *J Biol Chem* **269**: 16236-16241
- von Hippel PH, Pasman Z** (2002) Reaction pathways in transcript elongation. *Biophys Chem* **101-102**: 401-423
- Ward DF, Gottesman ME** (1981) The nus mutations affect transcription termination in *Escherichia coli*. *Nature* **292**: 212-215
- Weisberg RA, Gottesman ME** (1999) Processive antitermination. *J Bacteriol* **181**: 359-367
- Whalen W, Ghosh B und Das A** (1988) NusA protein is necessary and sufficient in vitro for phage lambda N gene product to suppress a rho-independent terminator placed downstream of nutL. *Proc Natl Acad Sci U S A* **85**: 2494-2498
- Wilson HR, Yu D, Peters HK,3rd, Zhou JG und Court DL** (2002) The global regulator RNase III modulates translation repression by the transcription elongation factor N. *EMBO J* **21**: 4154-4161
- Worbs M, Bourenkov GP, Bartunik HD, Huber R und Wahl MC** (2001) An extended RNA binding surface through arrayed S1 and KH domains in transcription factor NusA. *Mol Cell* **7**: 1177-1189
- Xia T, Becker HC, Wan C, Frankel A, Roberts RW und Zewail AH** (2003a) The RNA-protein complex: direct probing of the interfacial recognition dynamics and its correlation with biological functions. *Proc Natl Acad Sci U S A* **100**: 8119-8123
- Xia T, Frankel A, Takahashi TT, Ren J und Roberts RW** (2003b) Context and conformation dictate function of a transcription antitermination switch. *Nat Struct Biol* **10**: 812-819
- Xia T, Wan C, Roberts RW und Zewail AH** (2005) RNA-protein recognition: single-residue ultrafast dynamical control of structural specificity and function. *Proc Natl Acad Sci U S A* **102**: 13013-13018
- Yarnell WS, Roberts JW** (1999) Mechanism of intrinsic transcription termination and antitermination. *Science* **284**: 611-615
- Zhang G, Campbell EA, Minakhin L, Richter C, Severinov K und Darst SA** (1999) Crystal structure of *Thermus aquaticus* core RNA polymerase at 3.3 Å resolution. *Cell* **98**: 811-824

**Zhou Y, Mah TF, Yu YT, Mogridge J, Olson ER, Greenblatt J und Friedman DI** (2001) Interactions of an Arg-rich region of transcription elongation protein NusA with NUT RNA: implications for the order of assembly of the lambda N antitermination complex in vivo. *J Mol Biol* **310**: 33-49

**Zou C, Fujita N, Igarashi K und Ishihama A** (1992) Mapping the cAMP receptor protein contact site on the alpha subunit of Escherichia coli RNA polymerase. *Mol Microbiol* **6**: 2599-2605

**Zuber M, Patterson TA und Court DL** (1987) Analysis of nutR, a site required for transcription antitermination in phage lambda. *Proc Natl Acad Sci U S A* **84**: 4514-4518

**Zuiderweg ER** (2002) Mapping protein-protein interactions in solution by NMR spectroscopy. *Biochemistry* **41**: 1-7





## 6. Publikationsliste

### 6.1. Einzelarbeit A

Stefan Prasch, Sabine Schwarz, Anke Eisenmann, Birgitta M. Wöhl, Kristian Schweimer, und Paul Rösch (2006): Interaction of the Intrinsically Unstructured Phage  $\lambda$  N Protein with *Escherichia coli* NusA. *Biochemistry* **45**, 4542-4549

Ich habe die Domänen AR1 und AR2 einzeln kloniert und die jeweiligen Reinigungsprotokolle etabliert. Das  $\lambda$  N Konstrukt wurde von Sabine Schwarz kloniert und die Reinigung wurde von mir optimiert. Alle NMR-Experimente wurden von mir mit Hilfe von Kristian Schweimer gemessen und ausgewertet. Die Zuordnung der ebenso verwendeten carboxyterminalen Domäne von NusA wurde von Anke Eisenmann durchgeführt. Die Zuordnung der einzeln klonierten Domänen erfolgte durch mich, basierend auf den Daten von Anke Eisenmann. Das Manuskript wurde von Paul Rösch, Birgitta M. Wöhl, Kristian Schweimer, Anke Eisenmann und mir verfasst.

### 6.2. Einzelarbeit B

Anke Eisenmann, Sabine Schwarz, Stefan Prasch, Kristian Schweimer und Paul Rösch (2005): The *E. coli* NusA carboxy-terminal domains are structurally similar and show specific RNAP and  $\lambda$  N interaction. *Protein Science* **14**, 2018-2029

Die Arbeit wurde von Anke Eisenmann im Rahmen ihrer Dissertation durchgeführt. Es wurde von mir die Reinigung der verwendeten Proteine  $\alpha$ CTD und  $\lambda$  N etabliert. Die Titrations wurden von Anke Eisenmann und mir durchgeführt. Das Manuskript wurde von Paul Rösch, Anke Eisenmann, Kristian Schweimer und mir verfasst.

### 6.3. Einzelarbeit C

Stefan Prasch, Kristian Schweimer und Paul Rösch (2007): Structural Basis of Transcription Elongation Control: The NusA- $\alpha$ CTD Complex. In Bearbeitung

Alle aufgeführten Experimente wurden von mir geplant und durchgeführt. Die Reinigung der notwendigen Proteine wurde von mir etabliert. Die NMR-Experimente wurden von mir mit Hilfe von Kristian Schweimer geplant und gemessen. Ebenso wurde von mir und Kristian Schweimer die Struktur berechnet. Das Manuskript wurde von Paul Rösch und mir verfasst.

### 6.4. Einzelarbeit D

Stefan Prasch, Marcel Jurk, Robert S. Washburn, Max E. Gottesman, Birgitta M. Wöhrl und Paul Rösch (2007): *E. coli* NusA recognizes phage  $\lambda$  *nut spacer* RNA sequences. Under Revision, Journal of Biological Chemistry

Alle aufgeführten Experimente wurden von mir geplant und durchgeführt. Das NusA Konstrukt wurde von mir kloniert und die entsprechende Reinigung von mir etabliert mit technischer Unterstützung von Marcel Jurk. Alle Fluoreszenzexperimente wurden von mir geplant und durchgeführt. *In vivo* Experimente wurden von Robert S. Washburn und Max E. Gottesman mit meiner Mithilfe geplant und teilweise von mir im Rahmen eines Forschungsaufenthaltes an der Columbia Universität durchgeführt. Das Manuskript wurde von Max E. Gottesman, Birgitta M. Wöhrl, Paul Rösch und mir verfasst.

## Einzelarbeit A

Stefan Prasch, Sabine Schwarz, Anke Eisenmann, Birgitta M. Wöhrl, Kristian Schweimer, und Paul Rösch

Interaction of the Intrinsically Unstructured Phage  $\lambda$  N Protein with *Escherichia coli* NusA. *Biochemistry* **45**, 4542-4549 (2006)



## Interaction of the Intrinsically Unstructured Phage $\lambda$ N Protein with *Escherichia coli* NusA

Stefan Prasch, Sabine Schwarz, Anke Eisenmann, Birgitta M. Wöhrle, Kristian Schweimer,\* and Paul Rösch

Lehrstuhl Biopolymere, Universität Bayreuth, Universitätsstrasse 30, 95440 Bayreuth, Germany

Received November 16, 2005; Revised Manuscript Received February 15, 2006

**ABSTRACT:** N protein of the *Escherichia coli* phage  $\lambda$  ( $\lambda$ N) is involved in antitermination, a transcription regulatory process that is essential for the expression of delayed early genes during phage lytic development.  $\lambda$ N is an intrinsically unstructured protein that possesses three distinct binding sites interacting with the carboxy terminus of the *E. coli* host factor protein NusA, the viral *nutBoxB*-RNA, and RNA polymerase, respectively. Heteronuclear NMR experiments with  $\lambda$ N(1–53) in complex with NusA(339–495) revealed that upon complex formation the  $\lambda$ N-binding interface,  $\lambda$ N(34–47), adopts a rigid structure. NMR data also indicate the induction of a weak helical structure in the *nutboxB* RNA-binding region  $\lambda$ N(1–22) upon binding to NusA(339–495) even in the absence of RNA. Titration experiments of the  $\lambda$ N(1–53)–*nutBoxB* RNA complex with NusA(339–495) revealed that the ternary complex can be described in terms of two structurally independent binary interactions. Furthermore, chemical-shift perturbation experiments with different NusA constructs and different  $\lambda$ N peptides showed that only NusA(353–416) is involved in  $\lambda$ N binding. We found that only one molecule of NusA(339–426) binds to one molecule of  $\lambda$ N(1–53). We also clarified the role of the  $\lambda$ N-binding region and could show that N41–R47 also binds to NusA(339–495). Furthermore, we observe that  $\lambda$ N(1–22) adopts a helical fold upon binding to NusA(339–495), in agreement with one of the theoretical models of  $\lambda$ N action.

One of the cornerstones of protein biochemistry is the structure–function paradigm, describing the relationship between the biological function of a protein and its three-dimensional structure (1). Recently, this point of view changed because of the increased analyses of intrinsically unstructured proteins (2). It is estimated that about 4.2% of all proteins in prokaryotes and 33.0% of all proteins in eukaryotes are intrinsically unstructured (3). These proteins are mainly involved in cell-cycle control and translational as well as transcriptional regulation (3). Transcriptional modulation of RNA polymerases (RNAPs) by intrinsically unstructured viral proteins is common in prokaryotes and eukaryotes; examples for such proteins are the  $\lambda$ N protein (4, 5), the phage HK022 Nun protein (6), and the human immunodeficiency virus (HIV)<sup>1</sup> Tat protein (7, 8). All of these proteins are unstructured in the free state and employ similar mechanisms to regulate transcription in the host cells (9).

Transcriptional control by the regulation of RNA elongation is widely used in the microbial world. In the bacteriophage  $\lambda$ -*Escherichia coli* (*E. coli*) system, protein N of

phage  $\lambda$  ( $\lambda$ N) 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  $\lambda$ N depends upon the recognition of an RNA element called *nut* (N utilization) on the nascent phage transcript. This RNA element consists of two parts, *nutBoxA* and *nutBoxB*. N and NusA together with the *nutBoxB* sequence are sufficient to induce effective antitermination in vitro at many terminators located within a few hundred base pairs downstream of a *nutBoxB* site (10–13). Highly efficient, processive N-mediated antitermination also requires *E. coli* transcription elongation factors NusA, NusB, NusG, and NusE (S10), as well as *nutboxA* (12–15). 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 (16, 17).

The key component of the antitermination complex is the highly basic 107 amino acid  $\lambda$ N protein, which is largely unfolded in solution (4, 5).  $\lambda$ N consists of three functionally distinct regions with different interaction partners: M1–N22 binds the *nutBoxB* RNA, N34–R47 binds the carboxy-terminal part of *E. coli* NusA, R73–I107 forms the RNA polymerase (RNAP) binding region (5, 18).

The region M1–N22 of  $\lambda$ N, called the arginine-rich motif (ARM), binds to the *nutBoxB* RNA hairpin with an affinity of  $\sim 10^{-9}$  M (19–21). The structure of the ARM region and *nutBoxB* RNA has already been solved by nuclear magnetic resonance (NMR) spectroscopy. Upon RNA binding, the ARM region of  $\lambda$ N adopts a stable  $\alpha$ -helical conformation (22, 23). The complex formation enforces a GNRA-type

\* To whom correspondence should be addressed: Lehrstuhl Biopolymere, Universität Bayreuth, Universitätsstrasse 30, 95440 Bayreuth, Germany. Telephone: +49-921-553543. Fax: +49-921-553544. E-mail: kristian.schweimer@uni-bayreuth.de.

<sup>1</sup> Abbreviations: CD, circular dichroism; NMR, nuclear magnetic resonance;  $\lambda$ N, protein N of phage  $\lambda$ ; HhH, helix hairpin helix; ARM, arginine-rich motif; AR, acidic repeat; AR1, NusA(353–416); AR2, NusA(431–490); *E. coli*, *Escherichia coli*; HIV, human immunodeficiency virus; YUH, yeast ubiquitin hydrolase; HSQC, heteronuclear single-quantum coherence; NOE, nuclear Overhauser effect; NOESY, NOE spectroscopy.

tetraloop fold of the five nucleotide *nutBoxB* RNA loop with one base extruded. W18 then stacks on the top of the tetraloop fold and extends the RNA  $\pi$  stack by one residue. Randomized mutagenesis studies revealed that the complex interface is the principal determinant of functional antitermination (24).

The region N34–R47 interacts with the carboxy-terminal part of *E. coli* NusA (5). N34–R47 binds to NusA with an affinity of 70 nM (21, 24), nearly 2 orders of magnitude weaker than the ARM region binds to *nutBoxB* RNA. The NMR structure of the carboxy terminus of NusA (NusA339–495) in the free state showed that *E. coli* NusA(339–495) consists of two separate domains acidic repeat (AR)1, NusA(353–416), and AR2, NusA(431–490) (25), with either one adopting a compact helix–hairpin–helix (HhH)<sub>2</sub> fold (26). The two domains are connected by a 14-residue flexible linker region (26). The three-dimensional structure of the complex of  $\lambda$ N and the carboxy-terminal part of *E. coli* NusA has been solved by X-ray crystallography (27). It shows only interactions between AR1 and N34–L40. The second part of the  $\lambda$ N-binding site, N41–R47, could not be detected in the electron-density map. Furthermore, the X-ray data reported a 2:1 stoichiometry of AR1/ $\lambda$ N(34–47) (27). Isothermal titration calorimetry (ITC) and mutational studies suggested, however, that only the 1:1 complex between NusA and  $\lambda$ N is biologically relevant (27). On the basis of these results, it was suggested that, in full-length NusA, N41–R47 binds to AR2 (27).

The sequence of steps leading to the formation of the complex of  $\lambda$ N, NusA, and RNAP is still unclear. One model proposes that NusA binds to RNAP after the initiation process. Immediately after the *nutBoxB* RNA leaves the RNAP, N binds to the RNA tetraloop and is recruited to the transcription elongation complex. Complex formation transforms the RNAP into a termination-resistant transcription complex (5, 16, 28–30). In an alternative model,  $\lambda$ N, NusA, and RNAP associate prior to the exit of RNAP from the *nut* site.  $\lambda$ N is then able to scan the nascent RNA for the *nutBoxB* signal and finally binds to it, forming a processive antitermination complex (24).

In this work, we studied the complex of intact NusA(339–495) with  $\lambda$ N(1–53) by multidimensional, heteronuclear nuclear magnetic resonance (NMR) spectroscopy to clarify the role of AR1 and AR2 and to further investigate details of the  $\lambda$ N–NusA interaction as an essential building block of a functional antitermination complex.

## MATERIALS AND METHODS

**Expression and Purification of  $\lambda$ N(1–53) and NusA(339–495).** A synthetic gene containing the  $\lambda$ N(1–53) DNA sequence adapted to the *E. coli* codon usage was cloned into the expression vector pTKK 19 (31). *E. coli* strain BL21 DE3 (Novagen) was transformed with the resulting plasmid. The expressed protein consisted of an amino-terminal deca-His-tag fused to ubiquitin from *Saccharomyces cerevisiae* followed by  $\lambda$ N(1–53). For double labeling, cells were grown in M9 minimal medium supplemented with <sup>15</sup>NH<sub>4</sub>Cl and 0.2% <sup>13</sup>C D-glucose as the sole nitrogen and carbon sources, respectively (32).

The recombinant protein was purified by nickel-affinity chromatography (5 mL His-trap chelating column, GE Health

Care) under denaturing conditions, using a buffer containing 20 mM Tris/HCl at pH 7.95, 200 mM imidazole, 500 mM NaCl, and 6 M urea. After the protein was bound to the column, urea was removed by decreasing the urea concentration slowly over 20 column volumes. The protein was eluted with a buffer containing 20 mM Tris/HCl at pH 7.9, 200 mM imidazole, and 500 mM NaCl. Fractions containing the protein were dialyzed against 50 mM Tris/HCl at pH 7.9, 100 mM NaCl, and 1 mM  $\beta$ -mercaptoethanol. Ubiquitin was cleaved off using yeast ubiquitin hydrolase (YUH) (1 mg/mL) (31). The  $\lambda$ N(1–53) peptide was further purified by reversed-phase HPLC using a C-18 column (Macherey and Nagel, Germany) with a constant gradient of 80% acetonitrile/water and 0.1% TFA (v/v). The molecular mass of  $\lambda$ N(1–53) was checked by SDS–PAGE and ESI–MS.

A synthetically produced gene adapted to the *E. coli* codon usage harboring the AR1 domain of NusA(339–426) was cloned via the *Nde*I and *Bam*HI restriction sites into the *E. coli* expression vector pET19b (Novagen). The plasmid was transferred into the *E. coli* strain BL21 (DE3) (Novagen). The resulting soluble recombinant protein contained an N-terminal deca-His tag. For <sup>15</sup>N and <sup>13</sup>C labeling, cells were grown in M9 minimal medium with <sup>15</sup>NH<sub>4</sub>Cl and <sup>13</sup>C glucose as the sole nitrogen and carbon sources, respectively (32).

To purify NusA(339–426), bacterial cell pellets were lysed by sonication in 20 mM NaPO<sub>4</sub> at pH 7.4, 0.5 M NaCl, and 1 mM dithiothreitol (DTT). After centrifugation, the supernatant was loaded onto a Ni-ion-affinity column (His-trap chelating, GE Health Care) and eluted by applying an imidazole step gradient. Peak fractions containing NusA(339–426) were dialyzed against 50 mM Tris/HCl at pH 8.0, 150 mM NaCl, and 1 mM DTT. After dialysis, the N-terminal deca-His tag was cleaved off using PreScission protease (0.8 unit/100  $\mu$ g protein). The cleaved NusA(339–426) was dialyzed against 50 mM Tris/HCl at pH 7.4 and 1 mM DTT and further purified via a QXL column (GE Health Care) by a NaCl step gradient in the same buffer with up to 1 M NaCl. The eluted fractions containing NusA(339–426) were concentrated with Vivaspin concentrators (Vivascience, MWCO 5000 Da).

Unlabeled and <sup>15</sup>N-labeled NusA(339–495) was purified as described previously (26). Chemically synthesized  $\lambda$ N(34–47) peptide containing blocked endgroups was obtained from Coring (Gernsheim, Germany).

**NMR Sample Preparation.** NMR samples consisted of 0.5–1.5 mM  $\lambda$ N(1–53) in 10 mM KPO<sub>4</sub> at pH 6.8, 50 mM NaCl, 0.02% sodium azide, and 10% D<sub>2</sub>O. Titration of <sup>15</sup>N-labeled  $\lambda$ N(1–53) was done by gradually adding up to a 5-fold molar excess of unlabeled NusA(339–495). NMR experiments for characterizing  $\lambda$ N(1–53) in the NusA-bound state were performed in the presence of a 2-fold excess of NusA(339–495).

**NMR Spectroscopy.** NMR spectra were recorded on Bruker Avance 400 MHz, DRX 600 MHz, Avance 700 MHz, and Avance 800 MHz spectrometers with either room temperature or cryogenic-cooled triple-resonance probes equipped with pulsed field-gradient capabilities. The sample temperature was either 298 K for titration experiments or 308 K in the case of the assignment experiments for doubly labeled  $\lambda$ N(1–53) bound to NusA(339–495) to improve the NMR spectra. However, characteristic chemical-shift changes were independent of the temperature. For assignment of the  $\lambda$ N-

(1–53) resonances in the free and NusA(339–495)-bound states, standard double- and triple-resonance experiments were performed (33, 34). The steady-state  $\{^1\text{H}\}^{15}\text{N}$  nuclear Overhauser enhancement (NOE) experiment was recorded using water flipback pulses and a binomial 3-9-19 watergate for water suppression (33). A relaxation delay of 6 s was employed. In the saturated subspectrum, protons were saturated by applying a series of  $120^\circ$  pulses with a 5 ms interval during individual pulses for the final 3 s of the relaxation delay.  $\{^1\text{H}\}^{15}\text{N}$  steady-state NOEs were expressed as intensity ratios of the corresponding signals in the saturated and unsaturated subspectra. Amide resonances of NusA(339–426) in the free and  $\lambda\text{N}(1-53)$ -bound states were assigned by comparing the chemical shifts and NOE patterns in 3D  $^{15}\text{N}$ -edited NOE spectroscopy (NOESY)–heteronuclear single-quantum coherence (HSQC) experiments with the corresponding data of NusA(339–495).

The dissociation constant  $K_D$  was determined from the changes of chemical shifts of  $^{15}\text{N}$ -labeled  $\lambda\text{N}(1-53)$  observed in an  $^1\text{H},^{15}\text{N}$  HSQC after gradual addition of the corresponding unlabeled binding partner [NusA(339–495) or  $\lambda\text{N}$  peptides]. Changes of chemical shifts of signals in the fast-exchange limit were fitted to the following equation for a two-state model:

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

$\delta_{\text{obs}}$ ,  $\delta_{\text{p}}$ , and  $\delta_{\text{PL}}$  are the chemical shifts for the actual mixture, the free  $\lambda\text{N}(1-53)$  peptide, and the completely bound peptide, respectively.  $[P]_0$  is the total concentration of  $\lambda\text{N}(1-53)$ , and  $r$  describes the NusA(339–495)/ $\lambda\text{N}(1-53)$  ratio.

Normalized chemical-shift changes of the amide resonances were expressed as

$$\Delta\delta_{\text{norm}} = \sqrt{\Delta\delta(^1\text{H})^2 + (0.1\Delta\delta(^{15}\text{N}))^2}$$

$\Delta\delta(\text{X})$  represents the chemical-shift difference of spin X between free and bound states. Normalized chemical-shift changes larger than 0.04 ppm were considered as significant (35).

## RESULTS AND DISCUSSION

**Free  $\lambda\text{N}(1-53)$  Peptide and Its Interactions with *boxB* and NusA.** Multidimensional heteronuclear NMR spectroscopy yielded assignments for all  $^1\text{H}$ ,  $^{13}\text{C}$ , and  $^{15}\text{N}$  resonances of free  $\lambda\text{N}(1-53)$ , except for some arginine side chains in the ARM region  $\lambda\text{N}(1-22)$ . For these residues, severe spectral overlap made an unambiguous assignment impossible. In agreement with previous circular dichroism (CD) and NMR spectroscopic measurements (21), the narrow chemical-shift dispersion in  $^1\text{H},^{15}\text{N}$  HSQC (Figure 1A) as well as in  $^1\text{H},^{13}\text{C}$  HSQC spectra indicates structural disorder of the peptide (Figure 1A) (36, 37).  $\{^1\text{H}\}^{15}\text{N}$  steady-state NOE measurements at 14.1 T resulted in negative signal intensities or missing signals in the saturated subspectrum compared to the unsaturated subspectrum (Figure 1B). Thus, the intensity ratio of the saturated versus the nonsaturated

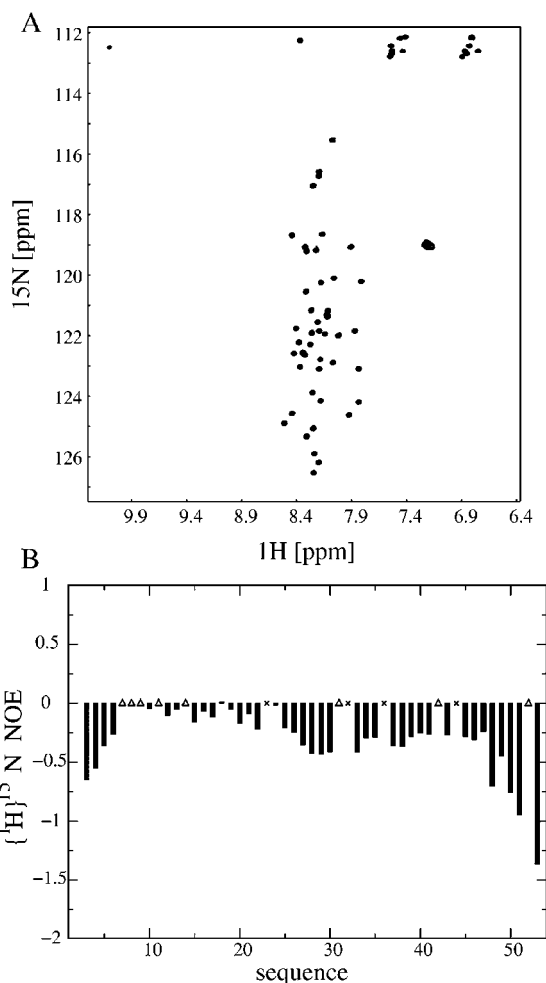


FIGURE 1: Free  $\lambda\text{N}(1-53)$  peptide.  $^1\text{H},^{15}\text{N}$  HSQC spectrum (800 MHz) (A) and  $\{^1\text{H}\}^{15}\text{N}$  NOE at 600 MHz of free  $\lambda\text{N}(1-53)$  peptide (B). The low dispersion together with the  $\{^1\text{H}\}^{15}\text{N}$  NOE values below 0 characterize the free  $\lambda\text{N}(1-53)$  peptide as an unstructured highly flexible peptide chain.

subspectrum is 0 or less for all residues, clearly pointing to a high flexibility of  $\lambda\text{N}(1-53)$ . Less negative values for the RNA-binding region, R10–A20, may be interpreted in terms of decreased flexibility of this region as compared to the rest of the peptide.

**Interaction Interfaces in the Ternary Complex Are Similar to That of the Binary Complexes.** The interaction between  $\lambda\text{N}(1-53)$ , *nutBoxB* RNA, and NusA(339–495) was monitored by observing chemical-shift changes of  $^{15}\text{N}$ -labeled  $\lambda\text{N}(1-53)$  after the gradual addition of unlabeled *nutBoxB* RNA and/or NusA(339–495) in  $^1\text{H},^{15}\text{N}$  HSQC spectra (Figure 2). On the basis of the titration experiments alone, it was not possible to completely assign the  $\lambda\text{N}(1-53)$  resonances of the respective complexes because some nuclei showed slow and intermediate exchange behavior during the titration. However, using multidimensional heteronuclear NMR spectra, all resonances of  $^{13}\text{C},^{15}\text{N}$ -labeled  $\lambda\text{N}(1-53)$  in the complex with NusA(339–495) could be assigned, with the exception of the backbone resonances of R11, R35, and P36 and a few side-chain resonances of longer amino acids.

Nearly all  $^1\text{H}$  and  $^{15}\text{N}$  resonances of  $\lambda\text{N}(1-53)$  change upon binding of the peptide to NusA(339–495), the most prominent changes occurring in the region I37–K45 (Figure 3A). This region has been implicated previously in complex

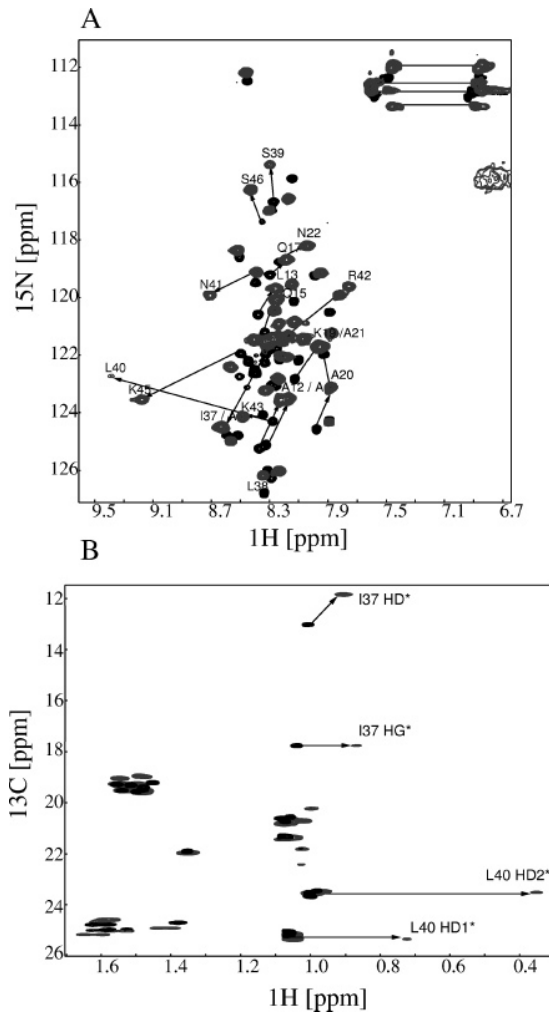


FIGURE 2: Interaction of  $\lambda$ N(1–53) and NusA(339–495). Overlay of  $^1\text{H}$ ,  $^{15}\text{N}$  HSQC (A) and methyl region of  $^1\text{H}$ ,  $^{13}\text{C}$  constant time HSQC (B) of free  $^{15}\text{N}$ ,  $^{13}\text{C}$ -labeled  $\lambda$ N(1–53) peptide (black) and in the presence of 2 equiv of unlabeled NusA(339–495) (gray). Resonances with remarkable changes of chemical shifts are labeled.

formation (5). Surprisingly, remarkable changes in the presence of NusA(339–495) could not only be detected within the expected binding region but also in the RNA-binding ARM region (residues M1–N22) of  $\lambda$ N(1–53).

To investigate whether changes in the ARM region could be observed upon the addition of NusA(339–495) in the presence of its known binding partner *nutBoxB* RNA, we gradually added unlabeled NusA(339–495) to a complex of  $^{15}\text{N}$ -labeled  $\lambda$ N(1–53) and *nutBoxB* RNA. Because changes could only be observed in the expected NusA-binding region, N34–R47, of  $\lambda$ N(1–53) (Figure 3B), we conclude that in the binary NusA(339–495)– $\lambda$ N(1–53) complex as well as in the ternary NusA(339–495)– $\lambda$ N(1–53)–*nutBoxB* RNA complex the conformation of residues N34–R47 of  $\lambda$ N(1–53) is virtually identical and thus is not influenced by *nutBoxB* RNA to a measurable extent.

**ARM Region of  $\lambda$ N(1–53) Possesses Helix Propensity That Can Be Enhanced by NusA(339–495).** Secondary chemical shifts of backbone  $^{13}\text{C}^\alpha$  and  $^{13}\text{C}'$  resonances are related to the peptide secondary structure (36). We referenced the chemical shifts of bound  $\lambda$ N(1–53) to the values of free  $\lambda$ N(1–53) rather than to published random-coil data, because the negative or close to 0  $\{^1\text{H}\}^{15}\text{N}$  steady-state NOEs at 14.1

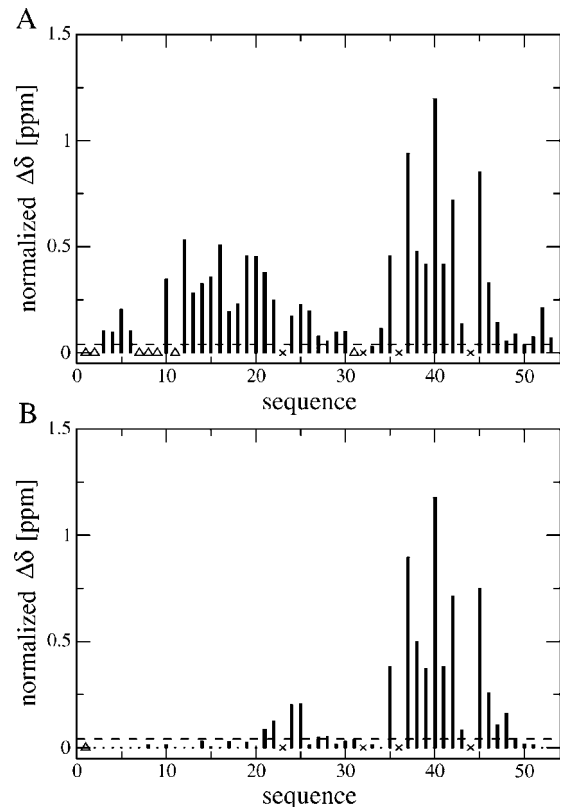


FIGURE 3: Backbone chemical-shift changes upon binding. (A) Investigation of the binary system. Changes of chemical shifts of  $\lambda$ N(1–53) after binding to NusA(339–495) as a function of the primary sequence. Normalized chemical-shift changes larger than 0.04 ppm were considered significant (35) ( $\times$ , proline;  $\Delta$ , missing residues). (B) Investigation of the ternary system. Differences of amide chemical shifts of  $\lambda$ N(1–53) complexed to *nutBoxB* RNA in the presence and absence of NusA(339–495). The lack of chemical-shift changes in the ARM region of  $\lambda$ N (M1–N22) complexed with *nutBoxB* RNA after the addition of NusA(339–495) indicates no ternary interaction, while the similar chemical-shift changes for residues I37–K45 point to the same structural binding mode between  $\lambda$ N(1–53) and NusA(339–495) in the presence or absence of *nutBoxB* RNA.

T of the free peptide indicate high flexibility of this peptide on the pico–nanosecond time scale and thus make it an ideal internal standard. Upon NusA(339–495) binding, notable changes for  $^{13}\text{C}^\alpha$  and  $^{13}\text{C}'$  can be observed in the amino-terminal RNA-binding region and around residues I37–K45 (Figure 4), in good agreement with changes of the  $^1\text{H}^{\text{N}}$  and  $^{15}\text{N}$  chemical shifts (Figure 3).

In particular, downfield shifts which are characteristic for helical secondary-structure formation are observed for  $^{13}\text{C}^\alpha$  and  $^{13}\text{C}'$  of the  $\lambda$ N ARM region after NusA(339–495) binding (Figure 4). The helical tendency is less pronounced for  $^{13}\text{C}^\alpha$  but can be clearly seen for the  $^{13}\text{C}'$  spins (Figure 4A). The chemical-shift changes extend over more than 13 amino acids, corresponding to three helical turns (Figure 4B). These changes indicate an increase of the population of the helical secondary structure for the RNA-binding region of  $\lambda$ N(1–53) in the NusA(339–495)– $\lambda$ N(1–53) complex even in the absence of *nutBoxB* RNA. Further evidence for the existence of an  $\alpha$  helix in the binary complex is the predominance of characteristic medium-range NOEs, such as  $[\text{d}\alpha\text{N}(i, i + 3)]$  and  $[\text{d}\alpha\beta(i, i + 3)]$  (38, 39). These NOE patterns could be found for at least six residues (E15, A16,



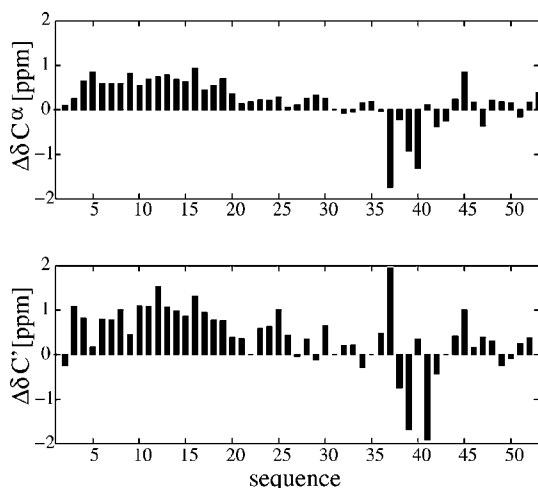


FIGURE 4:  $^{13}\text{C}^\alpha$  and  $^{13}\text{C}'$  chemical-shift changes upon binding. Changes of  $^{13}\text{C}^\alpha$  and  $^{13}\text{C}'$  chemical shifts of  $\lambda\text{N}(1-53)$  peptide after binding to NusA(339–495). Chemical-shift changes for  $^{13}\text{C}^\alpha$  and  $^{13}\text{C}'$  larger than 0.14 ppm were considered significant because of the spectral resolution. The changes of chemical shifts to a lower field for both nuclei indicates an increase of the propensity of the helical secondary structure for the ARM region after binding of  $\lambda\text{N}(1-53)$  to NusA(339–495) even in the absence of *nutBoxB* RNA.

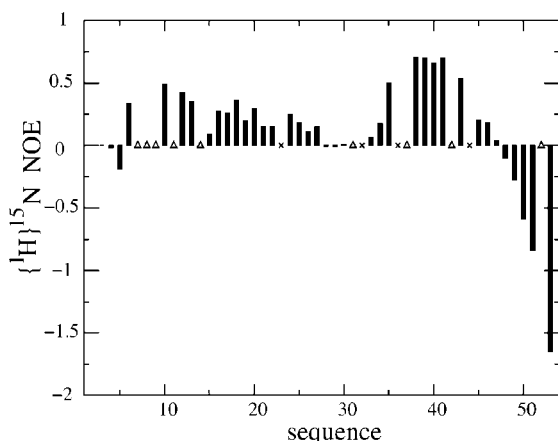


FIGURE 5:  $\{^1\text{H}\}^{15}\text{N}$  NOE in the bound state.  $\{^1\text{H}\}^{15}\text{N}$  steady-state NOE (600 MHz) of  $\lambda\text{N}(1-53)$  peptide in the presence of 2 equiv of NusA(339–395). Typical ranges ( $>0.65$ ) found in rigid-folded polypeptide chains are shown by the dashed line (40).

E17, W18, K19, and A20) in the carboxy-terminal part of the ARM motif.

Additionally, we measured the  $\{^1\text{H}\}^{15}\text{N}$  NOE at 14.1 T of  $\lambda\text{N}(1-53)$  in the NusA(339–495)– $\lambda\text{N}(1-53)$  complex (Figure 5). In comparison to the free state, positive NOE values can be found for the region T5–A20. This increase can mainly be attributed to the enhanced overall tumbling time after complex formation and reflects flexibility on the pico–nanosecond time scale in the bound state. However, the values did not reach the typical range ( $>0.65$ ) found in folded and rigid polypeptide chains (40). Therefore, a substantial flexibility on the pico–nanosecond time scale persists in the ARM region of  $\lambda\text{N}$  even after binding to NusA.

**N34–R47 of  $\lambda\text{N}(1-53)$  Binds to NusA(339–495) with a Well-Defined Structure.**  $\lambda\text{N}(1-53)$  binds to NusA(339–495) independently of *nutBoxB* RNA, and the relevant binding site includes residues N34–R47 (21, 24). Upon NusA(339–495) binding, the  $^1\text{H},^{15}\text{N}$  HSQC spectrum of  $\lambda\text{N}(1-53)$

exhibits significant changes of chemical shifts, e.g., L40 (1.2 ppm normalized chemical-shift change) and K45 (0.6 ppm). Furthermore, this region also exhibits the largest chemical-shift changes for  $^{13}\text{C}^\alpha$  and  $^{13}\text{C}'$  resonances (Figure 4) and considerable shifts of the methyl groups to a higher field (Figure 2B). High-field shifts of methyl groups are typically observed in hydrophobic cores of proteins or at binding interfaces (38, 39) and indicate a structurally defined complex formation. The chemical-shift changes for N34–L40 are thus consistent with results from X-ray crystallography, showing that these residues are involved in complex formation. Particularly, analysis of the crystal structure suggests that the upfield shifts of the L40 methyl resonances are due to a ring current effect caused by F369 of NusA (27).

In addition, the resonances of residues N41–K45 show chemical-shift changes, most prominent in K45 ( $^1\text{H}^\text{N}$ , 0.74 ppm), suggesting the involvement of these residues in complex formation, in agreement with data obtained with different deletion mutants of  $\lambda\text{N}$  (5). In that study, the peptide M1–R47 shows stronger binding affinity to NusA(1–495) than M1–P44. This finding together with the large chemical-shift change of K45 in our analyses implies an interaction of N41–K45 with the carboxy-terminal region of NusA, an interaction that could not be observed in the X-ray crystallographic studies because no electron density could be observed for N41–K45.

In contrast to A3–N22, the region N34–R47 shows no typical chemical-shift pattern for a regular secondary structure, in good agreement with the crystal structure, where N(34–40) adopts an extended conformation. The steady-state  $\{^1\text{H}\}^{15}\text{N}$  NOE suggests the peptide backbone to be rather rigid for residues I37–K43 (Figure 5), with  $\{^1\text{H}\}^{15}\text{N}$  NOEs values in the range of 0.5–0.7, which characterize this region as the most rigid part of  $\lambda\text{N}(1-53)$  after NusA(339–495) binding. Indeed, these values are close to those typically found in structured regions of proteins, where an NOE cutoff value of 0.65 at 14.1 T is commonly used to distinguish isotropic tumbling and conformational exchange (40). This indicates the formation of a well-defined structure for residues I37–K43 in the complex with NusA(339–495). The gradual decrease of the  $\{^1\text{H}\}^{15}\text{N}$  NOE values toward the carboxy terminus associated with an increase of flexibility may explain the absence of these residues in the X-ray crystallographic electron-density map.

**$\lambda\text{N}(1-53)$  Interacts Only with AR1.** It was shown previously by X-ray crystallography that the peptide-spanning residues N34–R47 of  $\lambda\text{N}$  bind to AR1, and only residues N34–L40 are visible in the electron density (27). Although the AR2 domain was not present in the crystal, it was suggested that it interacts with residues N41–R47 of  $\lambda\text{N}$  in the intact system. This model was based on three observations. First, peptide N34–R47 of  $\lambda\text{N}$  has a 3-fold higher affinity to NusA than peptide K31–K43 of  $\lambda\text{N}$ . Second, the  $\lambda\text{N}$  sequences N34–L40 (NRPILSL) and N41–R47 (NRKPKSR) are highly similar. Third, NusA(339–495) consists of two structurally equivalent domains (27).

Our titration experiments using  $^{15}\text{N}$ -labeled  $\lambda\text{N}(1-53)$  together with NusA(339–495) also indicate an interaction of NusA and the  $\lambda\text{N}$  region N41–R47. From these results alone, however, we cannot conclusively state whether AR2 is directly involved in binding. To better define the binding region of NusA(339–495), additional titrations using  $^{15}\text{N}$ -

labeled NusA(339–495) and unlabeled  $\lambda$ N peptides were performed. To avoid unspecific charge–charge interactions between the negatively charged NusA and the positively charged ARM region of  $\lambda$ N(1–53), synthetic peptides spanning residues N34–R47 and N41–R47 of  $\lambda$ N were used. Up to a 2-fold excess of N34–R47, only resonances of  $^{15}\text{N}$ -labeled AR1 residues changed their position in the  $^1\text{H}, ^{15}\text{N}$  HSQC spectra (Figure 6A), clearly demonstrating that only AR1 is involved in high-affinity binding to N34–R47 of  $\lambda$ N. Chemical-shift changes in AR2 could only be observed after saturation of AR1 with  $\lambda$ N(34–47), pointing to an unspecific interaction between positively charged N34–R47 of  $\lambda$ N (+5) and negatively charged AR2 (–9). In a second experiment, we titrated the peptide  $\lambda$ N(41–47) up to a 5-fold excess to  $^{15}\text{N}$ -labeled NusA(339–495). With the exception of residues V372, G377, S379, and L398, all located in AR1 (Figure 6B), no significant changes were observed.

The chemical-shift changes of amide resonances of NusA(339–495) during the titration with  $\lambda$ N(34–47) fit well to a two-state model (Figure 6C). The intersection of the two straight lines in the titration profile of  $^{15}\text{N}$ -labeled NusA(339–495) indicates a 1:1 ratio of NusA(339–495) and  $\lambda$ N(34–47), implying that a single NusA(339–495) molecule binds to one  $\lambda$ N(34–47) molecule (Figure 6C). The  $K_D$  value estimated for the peptide complex from the titrations was in the range of 3.7 to 5.0  $\mu\text{M}$ , comparable to values measured with ITC [3.55  $\mu\text{M}$  (27)].

Titrations of  $^{15}\text{N}$ -labeled  $\lambda$ N(1–53) with NusA(339–495) allowed for the observation of chemical-shift changes with fast-exchange behavior for resonances located in the RNA-binding region of  $\lambda$ N, and a  $K_D$  of 10–20  $\mu\text{M}$  could be estimated. Although the resonances used for  $K_D$  determination lie outside the binding region for NusA(339–495), the observed  $K_D$  is similar and therefore confirms the former measurement. The chemical-shift changes of residues located in the ARM region of  $\lambda$ N(1–53) after NusA binding were attributed to an increase of helical secondary structure, with the  $K_D$  value therefore reflecting the induction of the secondary structure outside the binding region rather than the direct interaction. Because both measured  $K_D$  values are nearly identical, the structural transition of the ARM region on NusA(339–495) binding correlates directly with the binding. Our NMR data indicate that residues N34–R47 are involved in binding to NusA(339–495), whereas residues N41–R47 probably do not contribute much to the binding energy. Differences in the  $K_D$  values of N34–R47 and K31–K43 are negligible, suggesting that the region N41–R47 contributes little to the binding energy (27). The differences in the  $K_D$  values measured with ITC and NMR, which are both in the same range, compared to values measured with fluorescence can be due to different salt conditions (21, 24, 27). Moreover, the high protein concentration required for the NMR experiments prohibits accurate  $K_D$  measurements in the low micromolar range, and the resulting  $K_D$  values reflect an upper limit only.

These titration experiments clearly demonstrate that all chemical-shift changes observed for NusA(339–495) are located in the region of AR1, strongly suggesting that only AR1 is the relevant part of the carboxy-terminal region of NusA for binding to  $\lambda$ N. Nevertheless, these experiments do not prove whether AR1 alone is responsible for all chemical-shift changes observed for  $\lambda$ N(1–53). To answer

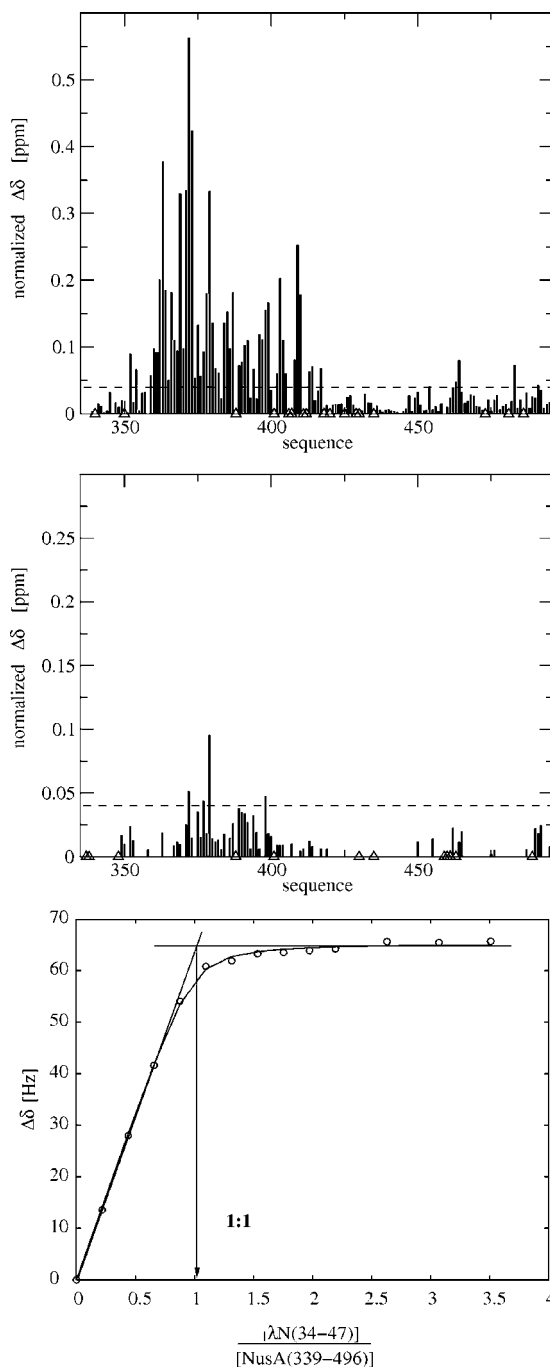


FIGURE 6: Titration of  $^{15}\text{N}$  NusA(339–495) with  $\lambda$ N peptides. Normalized chemical-shift changes of  $^{15}\text{N}$  NusA(339–495) in the presence of a 2-fold molar excess of unlabeled  $\lambda$ N(34–47) (A) or in the presence of a 3-fold molar excess of unlabeled  $\lambda$ N(41–47) (B). Normalized chemical-shift changes larger than 0.04 ppm (---) were considered significant (35). ( $\Delta$ , missing signals or overlap). The lack of chemical-shift changes in the AR2 domain of NusA(339–495) after the addition of  $\lambda$ N peptides indicate that AR2 is not involved in binding. (C) Stoichiometric titration of  $^{15}\text{N}$  NusA(339–495) with  $\lambda$ N(34–47) probed by NMR chemical-shift changes. The intersection at a ratio of 1:1 implies that one NusA(339–495) molecule binds to one  $\lambda$ N(34–47) molecule. L393 was used as an example. The determined  $K_D$  value is in the range of 9.8  $\mu\text{M}$ .

this question, NusA(339–426), containing AR1 only, was expressed, purified, and used for additional NMR experiments. Using  $^{15}\text{N}$ -labeled NusA(339–426) in a titration with unlabeled  $\lambda$ N(1–53), chemical-shift changes identical to

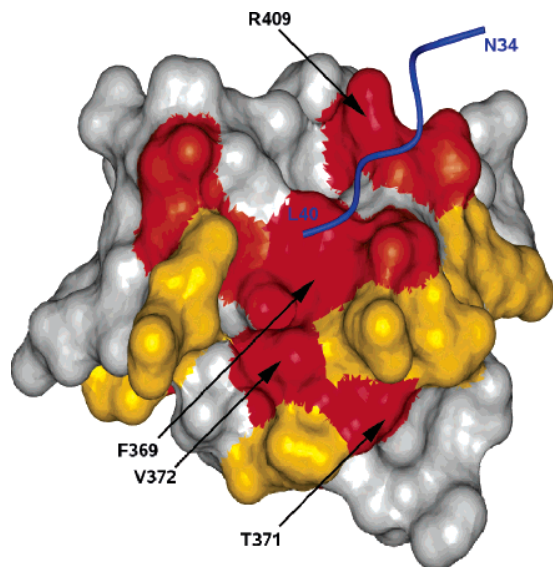


FIGURE 7: Mapping of the chemical-shift perturbations. Surface representation of the X-ray structure of NusA (339–426) in the presence of  $\lambda\text{N}(34\text{--}40)$  (27). The backbone of  $\lambda\text{N}(34\text{--}40)$  is shown in blue. The solvent-accessible surface of NusA (339–426) is shown in different colors according to the normalized chemical-shift perturbations ( $\Delta_{\text{norm}}$ ): gray,  $\Delta_{\text{norm}} < 0.09$  ppm; yellow,  $0.09 \text{ ppm} < \Delta_{\text{norm}} < 0.15$  ppm; orange,  $0.15 \text{ ppm} < \Delta_{\text{norm}} < 0.20$  ppm; and red,  $0.20 \text{ ppm} < \Delta_{\text{norm}}$ . Amino acids with the largest chemical-shift perturbations are indicated by arrows. The representation was produced with MOLMOL (42).

those already found for NusA(339–495) were observed (see Figure S1 in the Supporting Information). This unambiguously demonstrates that AR1 is the sole binding domain for  $\lambda\text{N}(1\text{--}53)$ . Additionally, the titration of  $^{15}\text{N}$ -labeled  $\lambda\text{N}(1\text{--}53)$  with unlabeled NusA(339–426) produced the same chemical-shift changes as the initial experiment with NusA(339–495). The prominent chemical-shift changes of the residues I37–K45 were present as well as the changes in the ARM region (see Figure S2 in the Supporting Information). It is clearly shown that AR1 alone causes the changes in the NMR spectrum of  $\lambda\text{N}(1\text{--}53)$  for the residues of the known binding region (N34–R47) and for the residues located in the ARM region.

Mapping the observed chemical-shift changes onto the surface of the crystal structure of the complex of AR1 with  $\lambda\text{N}(34\text{--}47)$  (Figure 7) gives further insight into the binding interface. Significant chemical-shift changes are found for the already known interface between AR1 and the region N34–L40. Additionally, large chemical-shift changes observed for F369, T371, V372, and L373 indicate the region of AR1 interacting with the additional residues of  $\lambda\text{N}$ , N41–R47. This region involves the carboxy-terminal part of helix 2 and the hairpin between helices 4 and 5.

Recent low-temperature studies of *nutBoxB* RNA showed a conformational equilibrium of this RNA between the conformation found in the  $\lambda\text{N}$ -bound state and other conformations (41). It is demonstrated that the amount of free RNA in the bound conformation is essential for binding, suggesting that enhancing the population of the binding conformation is a mechanism to facilitate complex formation. Short  $\lambda\text{N}$  peptides were described to populate the binding helical conformation at least to a small amount (19). Our structural NMR studies describe the increased population of

the helical conformation in the ARM region of  $\lambda\text{N}(1\text{--}53)$  after binding of  $\lambda\text{N}(1\text{--}53)$  to NusA AR1. In analogy to the situation described for *nutBoxB* RNA (41), it is likely that the preformation of the *nutBoxB*-bound  $\lambda\text{N}$  conformation in the ARM region on NusA-binding facilitates the recognition of the *nutBoxB* RNA. This could well be interpreted as lending additional support to the recently proposed model that  $\lambda\text{N}$  binds to NusA before the *nutBoxB* RNA interaction (24).

#### ACKNOWLEDGMENT

We thank U. Herzing and N. Herz for expert technical assistance. This project was funded by grants from the Deutsche Forschungsgemeinschaft (Ro617/8-4), the Sonderforschungsbereich 466, and the Bundesministerium für Bildung und Forschung.

#### SUPPORTING INFORMATION AVAILABLE

Overlay of  $^1\text{H}$ ,  $^{15}\text{N}$  HSQC spectra of  $^{15}\text{N}$ -labeled NusA-(339–426) in the absence and presence of  $\lambda\text{N}(1\text{--}53)$  (Figure S1) and overlay of  $^1\text{H}$ ,  $^{15}\text{N}$  HSQC spectra of  $^{15}\text{N}$ -labeled  $\lambda\text{N}(1\text{--}53)$  in the absence and presence of NusA(339–426) (Figure S2). This material is available free of charge via the Internet at <http://pubs.acs.org>.

#### REFERENCES

1. Uversky, V. N. (2002) Natively unfolded proteins: A point where biology waits for physics, *Protein Sci.* 11, 739–756.
2. Dyson, H. J., and Wright, P. E. (2005) Intrinsically unstructured proteins and their functions, *Nat. Rev. Mol. Cell. Biol.* 6, 197–208.
3. Fink, A. L. (2005) Natively unfolded proteins, *Curr. Opin. Struct. Biol.* 15, 35–41.
4. van Gilst, M. R., and von Hippel, P. H. (2000) Quantitative dissection of transcriptional control system: N-Dependent anti-termination complex of phage  $\lambda$  as regulatory paradigm, *Methods Enzymol.* 323, 1–31.
5. 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  $\lambda$  N protein, *Mol. Cell* 1, 265–275.
6. Faber, C., Scharpf, M., Becker, T., Sticht, H., and Rosch, 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.
7. Willbold, D., Rosin-Arbesfeld, R., Sticht, H., Frank, R., and Rosch, P. (1994) Structure of the equine infectious anemia virus Tat protein, *Science* 264, 1584–1587.
8. Bayer, P., Kraft, M., Ejchart, A., Westendorp, M., Frank, R., and Rosch, P. (1995) Structural studies of HIV-1 Tat protein, *J. Mol. Biol.* 247, 529–535.
9. Brigati, C., Giacca, M., Noonan, D. M., and Albin, A. (2003) HIV Tat, its targets and the control of viral gene expression, *FEMS Microbiol. Lett.* 220, 57–65.
10. Whalen, W., Ghosh, B., and Das, A. (1988) NusA protein is necessary and sufficient in vitro for phage  $\lambda$  N gene product to suppress a rho-independent terminator placed downstream of nutL, *Proc. Natl. Acad. Sci. U.S.A.* 85, 2494–2498.
11. Whalen, W. A., and Das, A. (1990) Action of an RNA site at a distance: Role of the nut genetic signal in transcription antitermination by phage- $\lambda$  N gene product, *New Biol.* 2, 975–991.
12. Mason, S. W., Li, J., and Greenblatt, J. (1992) Host factor requirements for processive antitermination of transcription and suppression of pausing by the N protein of bacteriophage  $\lambda$ , *J. Biol. Chem.* 267, 19418–19426.
13. DeVito, J., and Das, A. (1994) Control of transcription processivity in phage  $\lambda$ : Nus factors strengthen the termination-resistant state of RNA polymerase induced by N antiterminator, *Proc. Natl. Acad. Sci. U.S.A.* 91, 8660–8664.

14. Friedman, D. I., Olson, E. R., Johnson, L. L., Alessi, D., and Craven, M. G. (1990) Transcription-dependent competition for a host factor: The function and optimal sequence of the phage  $\lambda$  boxA transcription antitermination signal, *Genes Dev.* 4, 2210–2222.
15. Greenblatt, J., Nodwell, J. R., and Mason, S. W. (1993) Transcriptional antitermination, *Nature* 364, 401–406.
16. Mogridge, J., Mah, T. F., and Greenblatt, J. (1995) A protein–RNA interaction network facilitates the template-independent cooperative assembly on RNA polymerase of a stable antitermination complex containing the  $\lambda$  N protein, *Genes Dev.* 9, 2831–2845.
17. Friedman, D. I., Olson, E. R., Georgopoulos, C., Tilly, K., Herskowitz, I., and Banuett, F. (1984) Interactions of bacteriophage and host macromolecules in the growth of bacteriophage  $\lambda$ , *Microbiol. Rev.* 48, 299–325.
18. 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 Harbor Symp. Quant. Biol.* 63, 327–336.
19. Tan, R., and Frankel, A. D. (1995) Structural variety of arginine-rich RNA-binding peptides, *Proc. Natl. Acad. Sci. U.S.A.* 92, 5282–5286.
20. Cilley, C. D., and Williamson, J. R. (1997) Analysis of bacteriophage N protein and peptide binding to boxB RNA using polyacrylamide gel coelectrophoresis (PACE), *RNA* 3, 57–67.
21. van Gilst, M. R., and von Hippel, P. H. (1997) Assembly of the N-dependent antitermination complex of phage  $\lambda$ : NusA and RNA bind independently to different unfolded domains of the N protein, *J. Mol. Biol.* 274, 160–173.
22. Legault, P., Li, J., Mogridge, J., Kay, L. E., and Greenblatt, J. (1998) NMR structure of the bacteriophage  $\lambda$  N peptide/boxB RNA complex: Recognition of a GNRA fold by an arginine-rich motif, *Cell* 93, 289–299.
23. Scharpf, M., Sticht, H., Schweimer, K., Boehm, M., Hoffmann, S., and Rosch, P. (2000) Antitermination in bacteriophage  $\lambda$ . The structure of the N36 peptide–boxB RNA complex, *Eur. J. Biochem.* 267, 2397–2408.
24. Xia, T., Frankel, A., Takahashi, T. T., Ren, J., and Roberts, R. W. (2003) Context and conformation dictate function of a transcription antitermination switch, *Nat. Struct. Biol.* 10, 812–819.
25. 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.
26. Eisenmann, A., Schwarz, S., Prash, S., Schweimer, K., and Rosch, 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.
27. Bonin, I., Muhlberger, R., Bourenkov, G. P., Huber, R., Bacher, A., Richter, G., and Wahl, M. C. (2004) Structural basis for the interaction of *Escherichia coli* NusA with protein N of phage  $\lambda$ , *Proc. Natl. Acad. Sci. U.S.A.* 101, 13762–13767.
28. Chattopadhyay, S., Hung, S. C., Stuart, A. C., Palmer, A. G., III, Garcia-Mena, J., Das, A., and Gottesman, M. E. (1995) Interaction between the phage HK022 Nun protein and the nut RNA of phage  $\lambda$ , *Proc. Natl. Acad. Sci. U.S.A.* 92, 12131–12135.
29. Mason, S. W., and Greenblatt, J. (1991). Assembly of transcription elongation complexes containing the N protein of phage  $\lambda$  and the *Escherichia coli* elongation factors NusA, NusB, NusG, and S10, *Genes Dev.* 5, 1504–1512.
30. 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.
31. Kohno, T., Kusunoki, H., Sato, K., and Wakamatsu, K. (1998) A new general method for the biosynthesis of stable isotope-enriched peptides using a decahistidine-tagged ubiquitin fusion system: An application to the production of mastoparan-X uniformly enriched with  $^{15}\text{N}$  and  $^{13}\text{C}$ , *J. Biomol. NMR* 12, 109–121.
32. Marley, J., Lu, M., and Bracken, C. (2001) A method for efficient isotopic labeling of recombinant proteins, *J. Biomol. NMR* 20, 71–75.
33. Bax, A., and Grzesiek, S. (1993) Methodological advances in protein NMR, *Acc. Chem. Res.* 26, 131–138.
34. Schleucher, J., Schwendinger, M., Sattler, M., Schmidt, P., Schedletsky, 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.
35. 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.
36. Wishart, D. S., and Sykes, B. D. (1994) Chemical shifts as a tool for structure determination, *Methods Enzymol.* 239, 363–392.
37. Daughdrill, G. W., Hanely, L. J., and Dahlquist, F. W. (1998) The C-terminal half of the anti- $\sigma$  factor FlgM contains a dynamic equilibrium solution structure favoring helical conformations, *Biochemistry* 37, 1076–1082.
38. Wuthrich, K., Billeter, M., and Braun, W. (1984) Polypeptide secondary structure determination by nuclear magnetic resonance observation of short proton–proton distances, *J. Mol. Biol.* 180, 715–740.
39. Wagner, G., Neuhaus, D., Worgotter, E., Vasak, M., Kagi, J. H., and Wuthrich, K. (1986) Nuclear magnetic resonance identification of “half-turn” and 3(10)-helix secondary structure in rabbit liver metallothionein-2, *J. Mol. Biol.* 187, 131–135.
40. Pawley, N. H., Wang, C., Koide, S., and Nicholson, L. K. (2001) An improved method for distinguishing between anisotropic tumbling and chemical exchange in analysis of  $^{15}\text{N}$  relaxation parameters, *J. Biomol. NMR* 20, 149–165.
41. Johnson, N. P., Baase, W. A., and von Hippel, P. H. (2005) Low energy CD of RNA hairpin unveils a loop conformation required for  $\lambda$ N antitermination activity, *J. Biol. Chem.* 280, 32177–32183.
42. Koradi, R., Billeter, M., and Wuthrich, K. (1996) MOLMOL: A program for display and analysis of macromolecular structures, *J. Mol. Graphics* 14, 51–5, 29–32.

BI0523411

# **Interaction of the intrinsically unstructured phage $\lambda$ N protein with *E. coli* NusA\***

*Stefan Prasch, Sabine Schwarz, Anke Eisenmann, Birgitta M. Wöhrl, Kristian Schweimer,*

*and Paul Rösch*

From the Lehrstuhl Biopolymere, Universität Bayreuth, 95440 Bayreuth, Germany

## **Supporting information**

Corresponding address

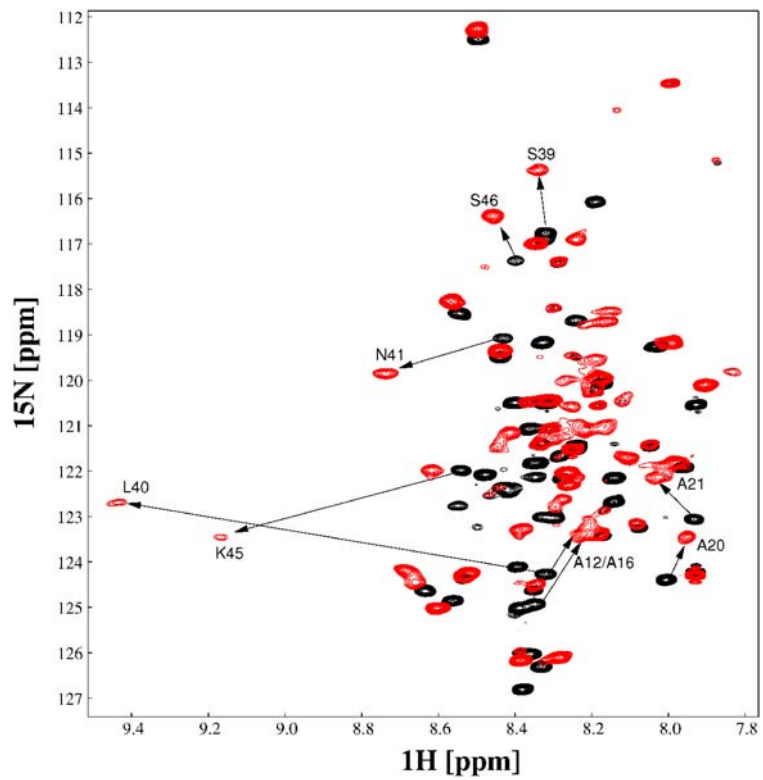
Dr. Kristian Schweimer, Lehrstuhl Biopolymere, Universität Bayreuth, 95440 Bayreuth, Germany;

Phone: +49 921 553543; Fax: +49 921 553544; kristian.schweimer@uni-bayreuth.de



Fig. S2: Interaction of NusA(339-426) and  $\lambda$ N(1-53)

Overlay of  $^1\text{H}$ ,  $^{15}\text{N}$  HSQC of free  $^{15}\text{N}$  labeled  $\lambda$ N(1-53) peptide (black) and in the presence of two equivalents of NusA(339-426) (red). Resonances with remarkable changes of chemical shifts are labeled.







## Einzelarbeit B

Anke Eisenmann, Sabine Schwarz, Stefan Prasch, Kristian Schweimer und Paul Rösch

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

*Protein Science* **14**, 2018-2029 (2005)



---

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

---

ANKE EISENMANN, SABINE SCHWARZ, STEFAN PRASCH, KRISTIAN SCHWEIMER, AND PAUL RÖSCH

Department of Biopolymers, University of Bayreuth, 95440 Bayreuth, Germany

(RECEIVED January 19, 2005; FINAL REVISION May 2, 2005; ACCEPTED May 4, 2005)

## Abstract

The carboxy-terminal domain of the transcription factor *Escherichia coli* NusA, NusACTD, interacts with the protein N of bacteriophage  $\lambda$ ,  $\lambda$ N, and the carboxyl terminus of the *E. coli* RNA polymerase  $\alpha$  subunit,  $\alpha$ CTD. We solved the solution structure of the unbound NusACTD with high-resolution nuclear magnetic resonance (NMR). Additionally, we investigated the binding sites of  $\lambda$ N and  $\alpha$ CTD on NusACTD using NMR titrations. The solution structure of NusACTD shows two structurally similar subdomains, NusA(353–416) and NusA(431–490), matching approximately two homologous acidic sequence repeats. Further characterization of NusACTD with  $^{15}\text{N}$  NMR relaxation data suggests that the interdomain region is only weakly structured and that the subdomains are not interacting. Both subdomains adopt an (HhH)<sub>2</sub> fold. These folds are normally involved in DNA–protein and protein–protein interactions. NMR titration experiments show clear differences of the interactions of these two domains with  $\alpha$ CTD and  $\lambda$ N, in spite of their structural similarity.

**Keywords:** NusA; anti-termination; termination; NMR; RNA polymerase; N-protein; phage  $\lambda$ ; HhH motif

RNA synthesis in *Escherichia coli* is catalyzed by RNA polymerase (RNAP), a multiprotein enzyme whose core shows an  $\alpha_2\beta\beta'$  subunit composition (Nudler 1999). After initiation of transcription, the essential transcription factor NusA (N utilization substance A) associates with the

RNAP core enzyme, where it modulates transcriptional pausing, termination, and anti-termination (Liu et al. 1996).

The crystal structures of two non-*E. coli* NusA factors have been solved so far (*Thermotoga maritima* [Worbs et al. 2001; Shin et al. 2003], *Mycobacterium tuberculosis* [Gopal et al. 2001]). These structures show a common domain organization, that is, an amino-terminal RNAP-binding domain, followed by one S1 and two KH (K homology) RNA-binding domains. This NusA core organization is conserved in all bacteria for which such sequence information is available. An additional carboxy-terminal region, NusACTD, comprising  $\sim 160$  residues is found in several  $\alpha$ -,  $\beta$ -, and  $\gamma$ -proteobacteria like the enterobacterium *E. coli*, as well as in *Chlamydia* or *Treponema* (Mah et al. 2000). Though NusACTD is not as highly conserved as the NusA core, the region is characterized by its acidity and frequently by an internal sequence repeat of  $\sim 70$  residues.

---

Reprint requests to: Paul Rösch, Department of Biopolymers, Universitätsstr. 30, University of Bayreuth, 95440 Bayreuth, Germany; e-mail: paul.roesch@uni-bayreuth.de; fax: +49-921-553544.

**Abbreviations:**  $\alpha$ CTD, carboxy-terminal domain of the  $\alpha$  subunit of the RNAP; *E. coli*, *Escherichia coli*; HetNOE,  $\{^1\text{H}\}^{15}\text{N}$  heteronuclear NOE; HhH, helix hairpin helix; HSQC, heteronuclear single quantum coherence; KH, K homology;  $\lambda$ N, protein N of phage  $\lambda$ ; NOESY, nuclear Overhauser spectroscopy; NusA, N utilization substance A; NusACTD, carboxy-terminal domain of NusA; NMR, nuclear magnetic resonance; PG-SLED, pulse gradient-stimulated echo longitudinal encode-decode; RDC, residual dipolar coupling; RMSD, root mean-square deviation; RNAP, RNA polymerase; SAM, sterile  $\alpha$  motif.

Article published online ahead of print. Article and publication date are at <http://www.proteinscience.org/cgi/doi/10.1110/ps.051372205>.

NusACTD probably serves as a multipurpose protein–protein interaction site as suggested by the existence of different binding partners of NusACTD in distinct elongation complexes (Mogridge et al. 1995, 1998; Mooney et al. 1998; Mah et al. 1999, 2000; Gusarov and Nudler 2001; Bonin et al. 2004). In spite of the fact that NusA harbors three RNA-binding domains, the affinity of isolated NusA to RNA is rather weak (Mogridge et al. 1995). Binding studies with deletion mutants of NusA, however, indicate that RNA binding to *E. coli* NusA is hindered by the 80 carboxy-terminal residues of NusACTD (Mah et al. 2000). Furthermore,  $\alpha$ CTD counteracts the inhibitory effect of the carboxy-terminal 80 residues by binding to NusACTD, thus providing a model for the regulation of RNA binding in the NusA:RNAP complex in vivo (Mah et al. 2000).

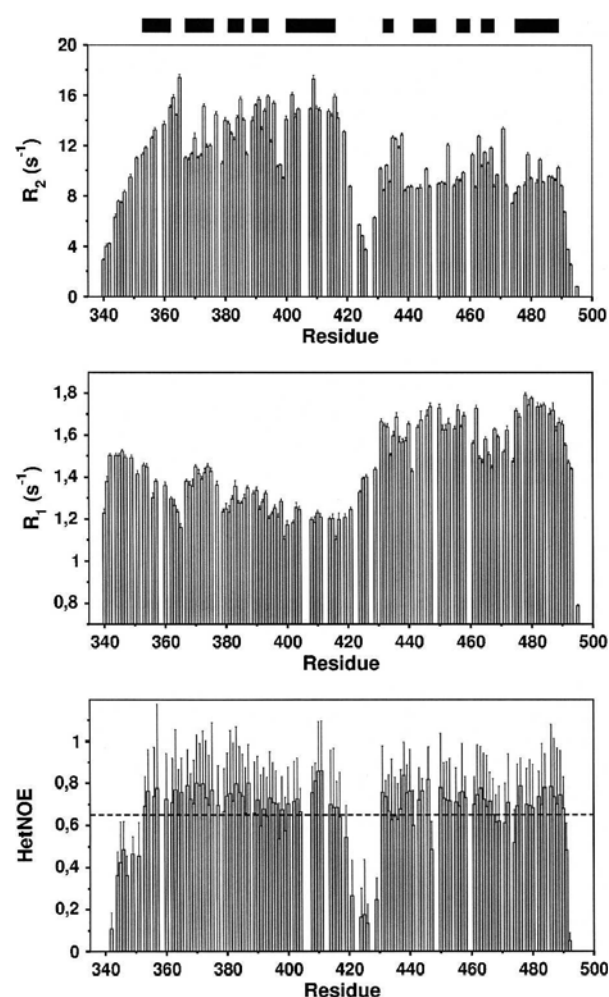
A similar mechanism to enhance the intrinsic RNA-binding capacity of NusA was suggested for the  $\lambda$ N protein (Mah et al. 2000), which binds to NusA during  $\lambda$ N-mediated anti-termination (for recent reviews, see Gusarov and Nudler 2001 and Nudler and Gottesman 2002). The minimal fragment of  $\lambda$ N required for the  $\lambda$ N–NusA interaction consists of residues N34–K47 (Mogridge et al. 1998). Recently, Bonin et al. (2004) determined the X-ray structure of the  $\lambda$ N(34–40)–NusA(352–421) complex and presented calorimetric data that indicated an interaction of  $\lambda$ N(41–47) with NusACTD outside NusA(352–421). However, in contrast to  $\alpha$ CTD,  $\lambda$ N(34–47) was not able to stimulate RNA binding to NusA (Mogridge et al. 1998; Bonin et al. 2004). Thus, Bonin et al. (2004) inferred a simple scaffold function for NusACTD in the  $\lambda$ N:NusA complex, and suggested different binding sites for  $\alpha$ CTD and  $\lambda$ N on NusA(431–490). A direct interaction between NusA(431–490) and  $\lambda$ N(41–47), however, has not been demonstrated yet, and it is not known whether  $\alpha$ CTD and  $\lambda$ N bind simultaneously to NusACTD in the anti-termination complex.

The modular architecture of the proteins establishing the termination and anti-termination complexes allows dissection and investigation of isolated domains and complexes thereof. We used this fact to determine the three-dimensional solution structure of NusA(339–495) to high resolution, its dynamic properties, and its interaction sites with  $\lambda$ N(1–53) and  $\alpha$ CTD(233–329) using NMR spectroscopy.

## Results and Discussion

### Structure calculation and validation

Initial analysis of  $^{15}\text{N}$  relaxation data (Fig. 1) revealed significant variations of the  $\{^1\text{H}\}^{15}\text{N}$  heteronuclear NOE (HetNOE) and the relaxation rates along the backbone.



**Figure 1.**  $^{15}\text{N}$  transverse ( $R_2$ , top) and longitudinal ( $R_1$ , middle) relaxation rates, and  $\{^1\text{H}\}^{15}\text{N}$  heteronuclear NOE (bottom) at 14.1 T. Negative values of the HetNOE for amino- and carboxy-terminal residues (T340, V341, D493, A495) have been omitted for clarity. The black bars represent helices. The evident difference of average  $R_2$  values for residues A350–A420 and residues P430–W490, together with the relatively low values ( $<0.65$ ; see dotted line) of the HetNOE in between residues A420 and P430 corroborates the existence of two subdomains separated by a mobile linker or hinge region. Quantitative analysis of the relaxation data indicates that the two subdomains reorient partially independent in solution and are noninteracting.

Values of the HetNOE  $<0.65$  at 600 MHz are indicative of a considerable flexibility on a picosecond timescale (Tjandra et al. 1995) and could be observed for the amino- and carboxy-terminal regions of NusACTD as well as residues I417–P430. Higher average values of the HetNOE  $\sim 0.7$  were found for NusA(353–416) and NusA(431–490), suggesting that these parts of the peptide chain are structured.

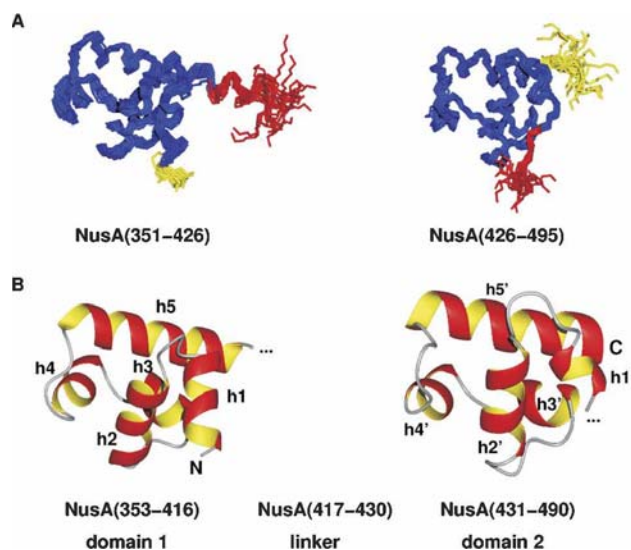
Both NusA(353–416) and NusA(431–490) exhibit similar average values of the longitudinal  $R_1$  rates,  $1.3 \pm 0.1 \text{ sec}^{-1}$

and  $1.6 \pm 0.9 \text{ sec}^{-1}$ , but show different average transversal  $R_2$  rates,  $13.6 \pm 1.9 \text{ sec}^{-1}$  and  $9.9 \pm 1.5 \text{ sec}^{-1}$ , respectively. As the  $R_2/R_1$  ratio for rigid residues is related to the rotational correlation time  $\tau_c$  (Kay et al. 1989), the different  $R_2/R_1$  ratios of NusA(353–416) and NusA(431–490), together with the low HetNOE values for NusA(417–430), suggest that NusACTD consists of two domains that are separated by a flexible linker region.

The hypothesis of NusACTD exhibiting a two-domain structure is supported by the fact that different ranges of residual dipolar couplings (RDCs) were observed for NusA(353–416),  $-12.4$ – $19.0$  Hz, and for NusA(431–490),  $-8.3$ – $11.0$  Hz, indicating different alignment tensors for these domains (Bax et al. 2001). In fact, in the course of molecular dynamics simulations, the magnitude of the alignment tensor  $D_a$  could be determined as  $10.5 \pm 0.25$  Hz [NusA(353–416)] and  $6.5 \pm 0.25$  Hz [NusA(431–490)], while the rhombicities  $R$  were  $0.2 \pm 0.05$  for NusA(353–416) and  $0.3 \pm 0.05$  for NusA(431–490). The average alignment of both domains cannot be described by a single alignment tensor; therefore, the significantly different tensor components (Table 1) imply considerable relative motion of NusA(353–416) and NusA(431–490) (Braddock et al. 2001, 2002).

All NOESY cross-peaks could be interpreted as short distances between protons within the domains NusA(353–416) and NusA(431–490), respectively, and no NOE between the two domains could be observed. Therefore, though NMR experiments were performed using the complete NusACTD, structures were calculated separately for NusA(351–426) and NusA(426–495). Residues 339–350 were not included in the structure calculation, as this region lacked nontrivial interresidual NOEs and showed low HetNOE values, indicating high flexibility of the amino terminus.

The structure calculation based on NMR-derived restraints resulted in ensembles with high precision of atomic coordinates for both domains (Fig. 2A). No violations of distance restraints  $> 0.16$  Å, dihedral angle restraints  $> 3^\circ$ , or RDCs  $> 0.72$  Hz could be observed (Table 2). Further validation of the structures with PROCHECK 3.5.4 (Morris et al. 1992; Laskowski et al. 1993) showed that 92.1% [NusA(351–426)] and 95.3% [NusA(426–495)] of the nonglycine and nonproline residues adopted a con-



**Figure 2.** Structure of the carboxy-terminal domain of NusA. (A) Backbone overlay of the 19 accepted structures for the two subdomains of the carboxy-terminal domain of NusA, NusA(351–426), and NusA(426–495), respectively. The linker residues are shown in red; flexible residues at the amino and carboxyl terminus are depicted in yellow. (B) Lowest energy structures of the relatively rigid parts of NusACTD, NusA(353–416), and NusA(431–490) in ribbon representation. Each subdomain contains two HhH motifs that are linked together by a so-called connector helix, thus adopting a compact (HhH)<sub>2</sub> fold. Helices h1/h1' and h2/h2' constitute the first and h4/h4' and h5/h5' the second HhH motif, whereas h3/h3' represent the connector helices. The linker between the two subdomains seems to be structured (see A, red part of overlay); however, <sup>15</sup>N relaxation data indicate that the interdomain region comprising about 14 amino acids is flexible. The picture was drawn with MOLMOL (Koradi et al. 1996).

formation within the most favored regions of the Ramachandran plot.

The solution structures of NusA(351–426) and NusA(426–495) were deposited in the Protein Data Bank (PDB) (Berman et al. 2000), accession codes 1WCL and 1WCN.

#### Structure of NusACTD

NusACTD consists of two subdomains, NusA(353–416) and NusA(431–490), that are connected by a linker region (Fig. 2B). Either subdomain contains two helix-hairpin-

**Table 1.** Alignment data for NusA(353–416) and NusA(431–490)

	$D_a$ (Hz)	$R$	$A_{zz}$ (Hz)	$A_{yy}$ (Hz)	$A_{xx}$ (Hz)
NusA(353–416)	$10.25 \pm 0.25$	$0.20 \pm 0.05$	$20.50 \pm 0.50$	$-13.65 \pm 1.11$	$-7.35 \pm 0.96$
NusA(431–490)	$6.25 \pm 0.25$	$0.30 \pm 0.05$	$12.50 \pm 0.50$	$-9.43 \pm 0.85$	$-3.85 \pm 0.63$

Shown are the magnitude of the alignment tensor  $D_a$  and the rhombicity  $R$ , as well as the eigenvalues of the alignment tensor  $A_{zz}$ ,  $A_{yy}$ , and  $A_{xx}$ . Different alignment tensors for individual domains indicate the presence of medium- to large-scale interdomain motions (Braddock et al. 2001).

**Table 2.** Summary of the structure calculation of NusA(351–426) and NusA(426–495)

Experimental restraints used for the structure calculation		NusA(351–426)	NusA(426–495)
Distance restraints	Total	883	699
	Sequential	328	285
	Medium-range	278	191
	Long-range	268	199
	Ambiguous/intramol.	9	24
Dihedral angles		39	28
Dipolar couplings		37	32
Hydrogen bonds (2 restraints each)		19	20
<i>Molecular dynamics statistics</i>			
Energies (kcal/mol)	$E_{\text{pot}}$	$23.1324 \pm 1.7314$	$17.4150 \pm 1.1643$
	$E_{\text{bond}}$	$1.0164 \pm 0.0879$	$0.8981 \pm 0.0962$
	$E_{\text{angle}}$	$8.7168 \pm 0.8408$	$6.5624 \pm 0.6648$
	$E_{\text{impr}}$	$2.6457 \pm 0.2801$	$1.6269 \pm 0.2001$
	$E_{\text{vdW}}$	$5.2270 \pm 1.0387$	$2.8016 \pm 0.6186$
	$E_{\text{NOE}}$	$4.0465 \pm 0.5038$	$3.4700 \pm 0.5815$
	$E_{\text{edih}}$	$0.0322 \pm 0.0495$	$0.0752 \pm 0.0846$
	$E_{\text{sani}}$	$1.4478 \pm 0.4740$	$1.9808 \pm 0.4104$
RMSDs from ideal distances (Å)	Bond lengths	$0.0007 \pm 0.0001$	$0.0009 \pm 0.0001$
	Distance restraints	$0.0054 \pm 0.0004$	$0.0097 \pm 0.0008$
RMSDs from ideal angles (deg)	Bond angles	$0.1169 \pm 0.0065$	$0.1520 \pm 0.0075$
	Dihedral angle restraints	$0.0655 \pm 0.0303$	$0.3179 \pm 0.2822$
RMSDs from dipolar couplings (Hz)		$0.1978 \pm 0.0327$	$0.2488 \pm 0.0249$
<i>Atomic RMSDs of structural ensemble (Å)</i>			
Backbone heavy atoms <sup>a</sup>		$0.34$ (H353–T416) <sup>b</sup>	$0.36$ (A431–W490) <sup>b</sup>
Heavy atoms <sup>a</sup>		$0.70$ (H353–T416) <sup>b</sup>	$0.72$ (A431–W490) <sup>b</sup>
<i>Ramachandran Plot statistics</i>			
Residues in	Most favored regions	92.1%	95.3%
	Allowed regions	7.0%	4.4%

Except for the experimental restraints and the atomic RMSD data, all values are average values over the 19 accepted structures in the format average value  $\pm$  standard deviation.

<sup>a</sup>Structural precision was calculated with reference to the structure with lowest value of the target function.

<sup>b</sup>All residues were included for calculation of the structural superposition.

helix (HhH) motifs, each formed by two anti-parallel  $\alpha$ -helices connected by a short hairpin (Doherty et al. 1996).

The amino-terminal HhH motif of NusA(353–416), HhH1, is composed of helices h1 (H353–Y362) and h2 (E367–E376) and interacts with the carboxy-terminal HhH motif of the first domain, HhH2, constituted by helices h4 (M389–E394) and h5 (E400–T416). Analogously, the amino-terminal HhH motif of NusA(431–490), HhH1', which comprises helices h1' (D432–L435) and h2' (R442–A449), packs on the carboxy-terminal HhH motif, HhH2', which includes helices h4' (I464–A468) and h5' (D475–C489).

The interhelical angles of the HhH motifs range from 119° to 129° (HhH1, 119°; HhH2, 124°; HhH1', 129°; HhH2', 122°), slightly less than the interhelical angles of 130°–155° described for HhH motifs earlier (Doherty et al. 1996; Shao and Grishin 2000). These deviations from canonical HhH

motifs can be interpreted in the context of the observed flexibility and the small database used for the determination of the interhelical angles in Doherty et al. (1996).

The tight packing of the two HhH motifs results in the formation of a compact hydrophobic core, which is completed by the helix connecting the two HhH motifs in each subdomain, helix h3 (L381–Y386) and helix h3' (L456–A460). This arrangement of two HhH motifs linked by a connector helix has been classified as a separate fold termed (HhH)<sub>2</sub> (Shao and Grishin 2000).

(HhH)<sub>2</sub> folds are mainly implicated in DNA binding (Shao and Grishin 2000), as confirmed by a search of the PDB (Berman et al. 2000) using the average structures of NusA(353–416) and NusA(431–490) with the DALI program (Holm and Sander 1996). However, both domains show similarity to the SAM (sterile  $\alpha$  motif) domain of

the ephrin 2 receptor (PDB 1b4f; Thanos et al. 1999). This is a typical representative of proteins whose (HhH)<sub>2</sub> fold mediates protein–protein interactions.

#### Comparison of the two subdomains of NusACTD

Overlay of the average structures of NusA(356–415) and NusA(431–490) results in a backbone RMSD of 1.9 Å, and thus confirms that the sequence homology of the acidic repeats is reflected in the three-dimensional structure. The superposition of the two domains (Fig. 3) shows similar positions and orientations for helices h2–h5 and h2'–h5', but not for helices h1 and h1', though all helices tend to be shorter in the second domain.

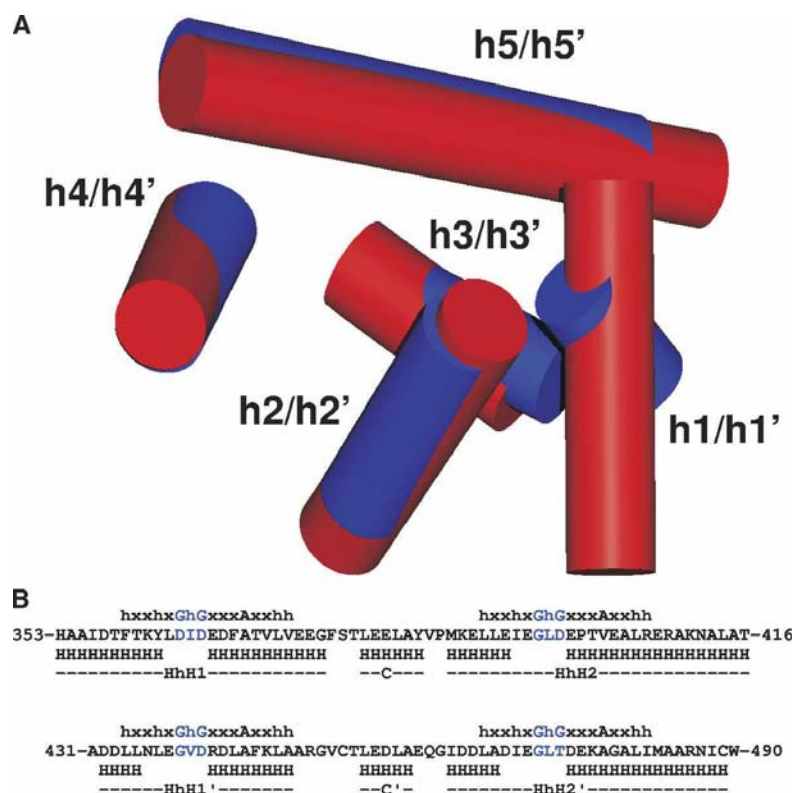
Both domains are considerably acidic with net charges of  $-11e^-$  and  $-9e^-$  and theoretical pIs of about 4. The

two domains show clear differences in the distribution of polar and hydrophobic as well as charged residues (Fig. 4), which may explain their differential recognition of binding partners.

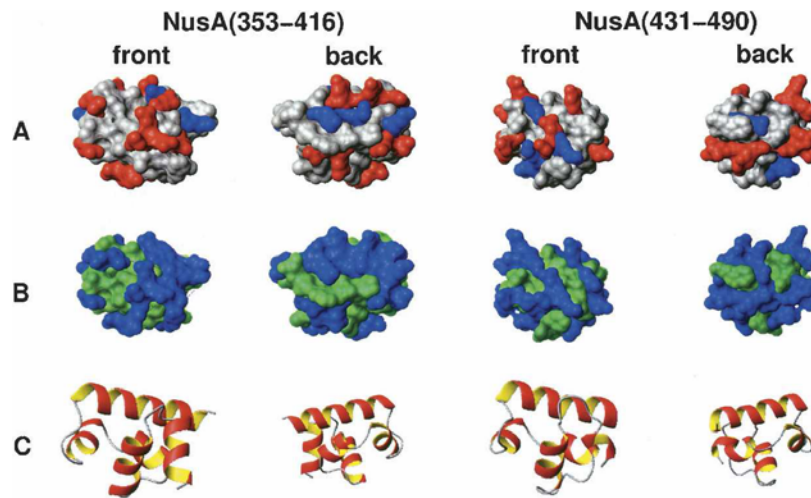
#### Oligomeric state of NusACTD

(HhH)<sub>2</sub> motifs have been implicated in oligomerization (Kim and Bowie 2003), which would have to be considered in the analysis of relaxation data.

To examine the possible formation of oligomers, we performed translational diffusion measurements (Wilkins et al. 1999) at concentrations of 0.175, 0.35, and 0.70 mM NusACTD. The resulting average hydrodynamic radius of  $r_H = 25.5 \pm 0.3$  Å was found to be independent of the protein concentration, which clearly indicates the virtual



**Figure 3.** The sequence homology of the two subdomains in the carboxy-terminal part of NusA is mirrored in the structural similarity of NusA(353–416) and NusA(431–490). (A) Overlay of NusA(353–416) and NusA(431–490). The arrangement of helices [NusA(353–416), h1–h5; NusA(431–490), h1'–h5'], which are represented as red and blue cylinders for NusA(353–416) and NusA(431–490), respectively, reveals a common overall fold. In contrast to helices h1 and h1', helices h2–h5 share similar positions and orientations with h2'–h5', though helices in the second subdomain tend to be shorter. The superposition was calculated for the backbone atoms of I356–A415 and A431–W490 and drawn using MOLMOL (Koradi et al. 1996). (B) Alignment of NusA(353–416) and NusA(431–490) with T-Coffee (Notredame et al. 2000) shows a sequence identity and homology of 31.3% and 60.9%, respectively. The secondary structure (H, helix), the helix–hairpin–helix (HhH) motifs as well as the connector helix (C) are reported beneath the sequence with primes for the second subdomain. For comparison, the 16-residue sequence characteristic of HhH motifs (Doherty et al. 1996) is shown *above* the sequence and the Gly–hydrophobic amino acid–Gly (GhG) pattern in the hairpin is highlighted in blue. Note that in all HhH motifs occurring in NusACTD, the second Gly that serves as a helix cap (Doherty et al. 1996) is replaced by more typical helix-capping residues like Asp or Thr (Kumar and Bansal 1998).



**Figure 4.** Surface representations of the lowest energy structures of NusA(353–416) and NusA(431–490) showing the distribution of charged (*A*) and polar residues (*B*). As a reference, the corresponding orientations of the subdomains are drawn in ribbon style (*C*). Both charged (negative, red; positive, blue; neutral, gray) and polar (polar, blue; hydrophobic, green) residues are differently distributed on the surface. The preparation of the figure was done using MOLMOL (Koradi et al. 1996).

absence of a fast monomer–oligomer equilibrium in the protein solution. In addition, the value of  $r_H$  cannot be explained in terms of dimer or higher oligomer formation of NusACTD. Although the hydrodynamic radius of a hypothetical spherical NusACTD can be estimated to be 20.7 Å for a monomer and 25.3 Å for a dimer (Wilkins et al. 1999), a more realistic calculation with an elongated NusA(353–490) structure using a shell model (de la Torre and Carrasco 2000) leads to a hydrodynamic radius of 22.7 Å. The difference between this value and the experimental value of 25.5 Å may easily be explained by the fact that the flexible amino- and carboxy-terminal residues could not be taken into account in this model. Thus, the experimental value for the hydrodynamic radius is virtually identical to the one calculated from a monomer model.

#### Rotational diffusion of NusACTD

$^{15}\text{N}$  relaxation data of anisotropically tumbling proteins contain information about the orientation of each  $\text{N-H}^{\text{N}}$  vector relative to the overall diffusion frame of the molecule. Therefore, this data is frequently used to determine the relative orientation of individual domains in multidomain proteins (for review, see Fushman et al. 2004). This approach, however, assumes domains with fixed relative orientations. In all other cases, the diffusion of the domains will not only be governed by the diffusion of the particle as a whole, but will be influenced to a certain extent by the diffusional properties of the single domains. Thus, by calculating the diffusion tensor separately for each domain within a multidomain protein, it is pos-

sible to determine whether the domains are mutually independent.

This is clearly the case for NusA(353–416) and NusA(431–490). Quantitative analysis of the  $^{15}\text{N}$  data reveals significantly different diffusion tensors for these domains (Table 3), leading to the conclusion that their movements are highly uncorrelated.

The rotational correlation times were calculated from the  $^{15}\text{N}$  relaxation rates  $R_1$  and  $R_2$  that were determined for residues inside NusA(353–416) and NusA(431–490) in the intact NusACTD. The resulting correlation times were  $9.5 \pm 0.3$  nsec and  $7.0 \pm 0.2$  nsec, respectively, both values being higher than the roughly 4.5 nsec expected for globular, compact proteins of 64 and 60 residues (Maciejewski et al. 2000). This difference, however, may easily be explained by the fact that unstructured regions precede and follow the two domains of NusACTD, and that these domains are connected by a linker. The linker and the terminal extensions affect the rotational correlation times of the isolated domains in different ways. First, the linker restricts the relative movements of the two domains, which leads to an increase in the tumbling times of NusA(353–416) and NusA(431–490). Second, flexible termini can enhance the rotational correlation times significantly (Tjandra et al. 1995), which is reflected in the tumbling time of NusA(353–416) compared with NusA(431–495).

Taken together, the alignment and relaxation data strongly suggest that NusA(353–416) and NusA(431–490) are not interacting or are preferentially orientated relative to each other. This justifies our initial approach to treat the domains separately.



**Table 3.** Diffusion tensor analysis from  $R_2$  and  $R_1$  data at 600 MHz of NusA(353–416) and NusA(431–490)

Diffusion tensor statistics <sup>a</sup>								
	<i>Isotropic</i>		<i>Axially symmetric (prolate)</i>		<i>Axially symmetric (oblate)</i>		<i>Anisotropic</i>	
	$\chi^2$ (5%)	$\chi^2_{\text{exp}}$	$\chi^2$ (5%)	$\chi^2_{\text{exp}}$	$\chi^2$ (5%)	$\chi^2_{\text{exp}}$	$\chi^2$ (5%)	$\chi^2_{\text{exp}}$
NusA(353–416)	<b>44.2</b>	<b>207.6</b>	<b>40.0</b>	<b>48.5</b>	<b>39.7</b>	<b>48.5</b>	<b>37.0</b>	<b>19.0</b>
NusA(431–490)	31.4	55.98	26.6	24.3	28.0	39.7	23.4	22.9

Diffusion tensor components <sup>b</sup>						
	$D_{zz}$ ( $10^7$ s <sup>-1</sup> )	$D_{yy}$ ( $10^7$ s <sup>-1</sup> )	$D_{xx}$ ( $10^7$ s <sup>-1</sup> )	$\zeta$	$\eta$	$\tau_{c,\text{eff}}$ [ns]
NusA(353–416)	<b>2.31 ± 0.05</b>	<b>1.73 ± 0.06</b>	<b>1.22 ± 0.05</b>	1.6 ± 0.1	0.91 ± 0.19	9.5 ± 0.3
NusA(431–490)	3.11 ± 0.12	2.01 ± 0.07	2.01 ± 0.07	1.6 ± 0.1	0	7.0 ± 0.2

<sup>a</sup>  $\chi^2$  (5%) corresponds to the  $\alpha = 0.05$  confidence limit for the fit derived from 500 Monte Carlo simulations,  $\chi^2_{\text{exp}}$  refers to the value of the target function used for the fit. Values of  $\chi^2$  (5%) and  $\chi^2_{\text{exp}}$  for accepted models are shown in bold italic. An F-statistic was used to differentiate between the prolate (four parameters) and the fully asymmetric model (six parameters) for NusA(431–490). The critical values for the F-statistic with  $\alpha = 0.1$  amount to  $F_{\text{exp}} = 4.6$  and  $F_{(10\%)} = 22.3$ , indicating that the improvement of the fit upon using the six-parameter model is not statistically relevant (Mandel et al. 1995; Dosset et al. 2000). NusA(431–490) is thus best described by a prolate axially symmetric diffusion model, NusA(353–416) by an anisotropic diffusion tensor.

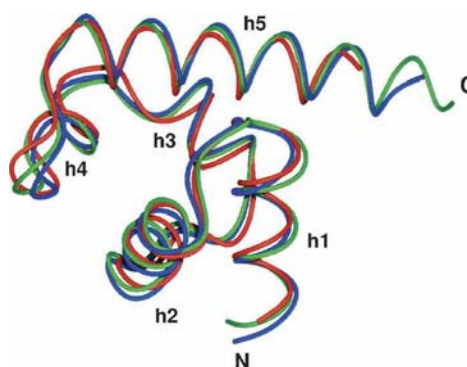
<sup>b</sup> Diffusion tensor components from  $R_2$  and  $R_1$  data at 600 MHz [NusA(353–416): 33 N-H<sup>N</sup> vectors; NusA(431–490): 21 N-H<sup>N</sup> vectors] for accepted models of NusA(353–416) and NusA(431–490). For a prolate diffusion tensor ( $D_{zz} - D_{yy} \geq D_{yy} - D_{xx}$ ) the anisotropy and rhombicity are defined as  $\zeta = 2D_{zz}/(D_{xx} + D_{yy})$  and  $\eta = 3/2 \cdot (D_{yy} - D_{xx})/D_{zz} \cdot \zeta/(\zeta - 1)$ , respectively, and characterize the deviations from a spherical top (Fushman et al. 2004). The effective rotational correlation time is related to the diffusion tensor components via  $\tau_{c,\text{eff}} = 1/(2D_{xx} + 2D_{yy} + 2D_{zz})$ . Parameter uncertainties of the diffusion parameters were taken from 500 Monte Carlo simulations.

### Comparison of free and complexed NusACTD

Recently, the structure of the complex of  $\lambda$ N(34–47) with two fragments of NusA, NusA(352–419), and NusA(352–421), respectively, has been determined by X-ray crystallography (Bonin et al. 2004). The NusA molecules are crystallographically independent, but structurally similar with an RMSD of the backbone heavy atoms of 0.67 Å for H353–T416. Both structures deviate only slightly from the uncomplexed solution structure of NusA(353–416), with an RMSD of the backbone heavy atoms of 1.20 Å and 1.10 Å (Fig. 5). Only minor differences between the free and the complexed structures could be observed, and most of these differences are located in the amino-terminal helix and the hairpin of the second HhH motif, involving M389–D399. These residues feature decreased HetNOE values in the free NusACTD, reflecting increased flexibility on a pico- to nanosecond time scale.  $\lambda$ N(34–47) contacts this region via L40 by side-chain–side-chain interactions with L398, supporting the idea that the flexibility of the helix and the adjacent hairpin might play an important role in the recognition (Atkinson and Kieffer 2004). Thus, in general, conformational rearrangements upon complex formation are minor, in line with the suggested function of NusA(353–416) as a scaffold conferring additional stability to the complex (Bonin et al. 2004).

### Interactions of NusACTD with $\alpha$ CTD and $\lambda$ N

In order to elucidate the sites of NusACTD that bind  $\lambda$ N and  $\alpha$ CTD, <sup>15</sup>N-labeled NusACTD was titrated



**Figure 5.** Overlay of the uncomplexed average structure of NusA(353–416) solved by NMR (this work, red) with the crystal structure 1U9L (Bonin et al. 2004), in which NusA(352–419) (chain A in 1U9L, blue) and NusA(352–421) (chain B in 1U9L, green) form a complex with  $\lambda$ N(34–47), respectively. Chains A and B are crystallographically independent, but almost identical to each other with an RMSD of the backbone heavy atoms of 0.67 over residues H353–T416. Both structures deviate only slightly from the uncomplexed NusA(353–416), with a backbone RMSD of 1.20 and 1.10, for chains A and B, respectively. The overlay was calculated and drawn with MOLMOL.

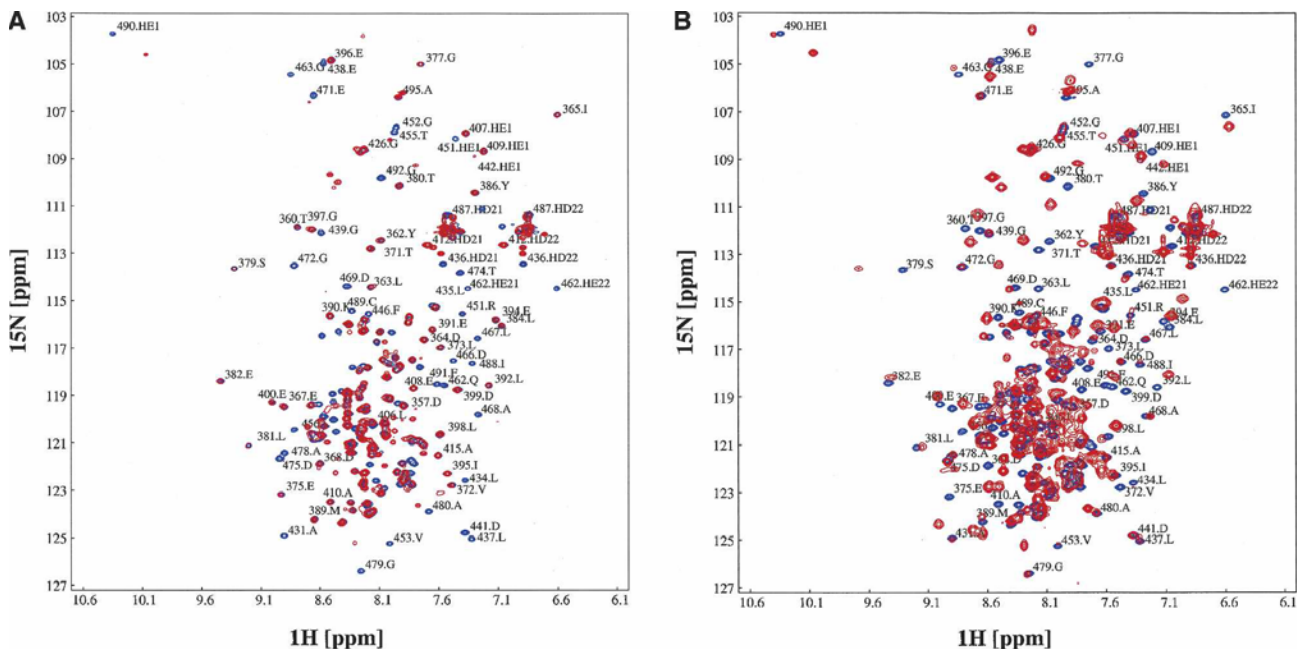
with unlabeled  $\alpha$ CTD or unlabeled  $\lambda$ N(1–53). HSQC spectra of the sample were taken after each titration step (Fig. 6).

Line broadening or chemical-shift changes could be observed for approximately two-thirds of the resonances of NusACTD throughout the titration with  $\lambda$ N(1–53), beginning with a NusACTD: $\lambda$ N(1–53) ratio of 0.4 and ending with a ratio of 3.4. Most resonances were broadened beyond detection starting from a NusACTD: $\lambda$ N(1–53) ratio of 1.3 and thus could not be tracked during the titration. From a ratio of 2.6, resonances appeared at new positions in the spectra, probably from the NusACTD: $\lambda$ N(1–53) complex. This broadening and re-appearance of resonances is typical for intermediate exchange processes (Zuiderweg 2002) and prevented detection of resonances for a wide range of concentration ratios.

During the titration, 98% and 70%, respectively, of the assigned backbone amide cross-peaks in the first and second domain showed significant changes in resonance intensity or position ( $\geq 0.05$  ppm in  $^1\text{H}$  or  $\geq 0.4$  ppm in  $^{15}\text{N}$  dimension). As most residues were affected by the presence of  $\lambda$ N(1–53), it was impossible to delineate a contact interface of  $\lambda$ N(1–53) on NusACTD. Widespread chemical-shift perturbations may indicate allo-

steric processes (Stevens et al. 2001). Comparison of the free and complexed NusA(353–416), however, clearly shows that the first domain of NusACTD does not undergo major rearrangements upon complex formation. Although no complex structure is available for the second domain of NusACTD,  $\lambda$ N is believed to bind in a similar fashion to both domains (Bonin et al. 2004), rendering large-scale structural changes unlikely. Instead, the scattered chemical-shift perturbations might be explained by subtle changes in the tertiary structure. Nevertheless, given that  $\lambda$ N(1–53) and NusACTD are strongly oppositely charged, unspecific interactions cannot be ruled out. In summary, both domains of NusACTD are affected by the presence of  $\lambda$ N(1–53), providing direct evidence for an interaction of  $\lambda$ N(1–53) with NusA(353–416) and NusA(431–490).

Addition of  $\alpha$ CTD to  $^{15}\text{N}$ -labeled NusACTD resulted in changes in HSQC resonance intensity or position for approximately one-third of the total NusA(339–495) backbone amide resonances from the first titration step (NusACTD: $\alpha$ CTD ratio of 0.5). Most of the perturbed resonances exhibited severe line broadening above a NusACTD: $\alpha$ CTD ratio of 1.1, and thus, could not be followed during the titration. Some of the affected resonances remained visible at all examined NusACTD:



**Figure 6.** Titration of NusACTD with  $\alpha$ CTD and  $\lambda$ N(1–53). (A) Interaction of NusACTD with  $\alpha$ CTD.  $^{15}\text{N}$ -labeled GP-NusA(339–495) was titrated with unlabeled  $\alpha$ CTD until no further changes were observed in the spectra, resulting in a NusACTD: $\alpha$ CTD ratio of 1:2.1. Shown are the overlays of the  $^1\text{H}$  $^{15}\text{N}$ -HSQC spectra at 700 MHz of the end points of the titrations, corresponding to the free (blue) and complexed (red) NusACTD. (B) Interaction of NusACTD with  $\lambda$ N(1–53).  $^{15}\text{N}$ -labeled GP-NusA(339–495) was titrated with  $\lambda$ N(1–53) until no further changes were observed in the spectra, resulting in a NusACTD: $\lambda$ N(1–53) ratio of 1:3.4. Shown are the overlays of the  $^1\text{H}$  $^{15}\text{N}$ -HSQC spectra at 700 MHz of the end points of the titrations, corresponding to the free (blue) and complexed (red) NusACTD.

$\alpha$ CTD ratios up to the end point of the titration (NusACTD: $\alpha$ CTD ratio of 2.1), again indicating intermediate exchange rather than protein aggregation.

In total, chemical-shift changes could be observed for 98% of the resolved backbone amide resonances of the second domain, respectively, whereas the resonances in the first domain remained unaffected by the presence of  $\alpha$ CTD. Though no defined contact interface of  $\alpha$ CTD on NusACTD could be derived by the titration experiments, the data clearly argue for  $\alpha$ CTD binding solely to NusA(431–490). Together with structural data and the fact that  $\alpha$ CTD is able to relieve the inhibitory effect of NusACTD on RNA binding (Mah et al. 2000), the titration data strongly suggest that NusA(431–490) functions as a domain regulating RNA binding.

#### *Implications for the role of NusACTD in termination and anti-termination*

In summary, the structural data provided here suggests how NusACTD functions as a versatile protein–protein interaction site in different transcription complexes. The domain structure offers an explanation of how NusA(431–490) regulates RNA binding to NusA. The flexible linker allows rearrangement of NusA(431–490) relative to the rest of the protein to facilitate access to the RNA-binding sites of NusA on complexation with  $\alpha$ CTD. Spatial adaptability might also play a role in switching from the termination-competent NusA:RNAP complex to the  $\lambda$ N:NusA:RNAP complex, which is capable of anti-termination.

In addition to its regulatory function, the NusA(431–490): $\alpha$ CTD interaction possibly enhances the stability of the NusA:RNAP complex (Liu et al. 1996). This could explain why a mutant NusA with residues L344–A495 missing prevents bacterial growth at elevated temperatures (Tsugawa et al. 1988).

We demonstrated in NMR-titration experiments that  $\lambda$ N(1–53) binds to both domains of NusACTD, NusA(353–416), and NusA(431–490). The region responsible for the NusACTD interaction on  $\lambda$ N is presumably located within residues N34–K47 (Mogridge et al. 1998). Despite its probable binding to NusA(431–490),  $\lambda$ N(34–47) cannot stimulate NusA's intrinsic RNA-binding capacity (Bonin et al. 2004). NusA, however, does bind sequence specifically to *nut* RNA in  $\lambda$ N:NusA:*nut* complexes (Mogridge et al. 1995; Mah et al. 2000). The  $K_D$  of full-length  $\lambda$ N and NusA or a  $\lambda$ N:*nutBoxB*:NusA complex is of the order of 50 nM (van Gilst and von Hippel 1997; Xia et al. 2003), suggesting that  $\lambda$ N or the  $\lambda$ N:*nutBoxB* complex might simply displace NusA(431–490) in the competition for the RNA-binding site. This is in accordance with the complex-stabilizing function of the NusA(353–416)– $\lambda$ N(34–40) interaction.

To get further insight into the role of NusA(353–416) and NusA(431–490) in termination and anti-termination, our future research will focus on the elucidation of the complexes of NusACTD with  $\lambda$ N and  $\alpha$ CTD, as well as investigation of the intramolecular interaction of NusA(431–490) and the RNA-binding sites of NusA.

## Materials and methods

### *Protein production*

$\lambda$ N(1–53) was expressed in *E. coli* BL21(DE3) using the pTKK19 expression vector (Kohno et al. 1998). The resulting fusion protein consisted of an amino-terminal deca-histidine-tag, followed by ubiquitin and  $\lambda$ N(1–53), and was purified by immobilized nickel-affinity chromatography under nonnative conditions. Subsequent cleavage by yeast ubiquitin hydrolase and a further high-performance liquid chromatography step yielded pure  $\lambda$ N(1–53). *E. coli*  $\alpha$ CTD(233–329) was expressed as a deca-histidine-tagged protein in BL21(DE3) from pET-19b (Novagen) and purified by immobilized nickel-affinity chromatography under native conditions. The histidine-tag was cleaved off with rhinovirus protease 3C (Cordingley et al. 1990), resulting in two additional amino-terminal residues (Gly-Pro). NusACTD was expressed and purified as GP-NusA(339–495) (Eisenmann et al. 2004). NMR samples contained 0.5–1.5 mM GP-NusA(339–495) in 10 mM potassium phosphate (pH 6.8), 50 mM sodium chloride, 0.02% sodium azide, and 10% D<sub>2</sub>O.

### *NMR spectroscopy*

All NMR experiments were performed at 298 K on Bruker Avance400, DRX600, Avance700, and Avance800 spectrometers equipped with inverse <sup>1</sup>H/<sup>13</sup>C/<sup>15</sup>N triple-resonance probes with pulsed-field gradient capabilities. In addition to the spectra needed for backbone and side-chain resonance assignments (Eisenmann et al. 2004), <sup>13</sup>C-NOESY-HSQC, <sup>15</sup>N-NOESY-HSQC, CNH-NOESY, CCH-NOESY, NNH-NOESY, and 4D-C,C-NOESY experiments with mixing times of 120 msec were acquired on uniformly <sup>15</sup>N and <sup>13</sup>C-labeled GP-NusA(339–495) at a concentration of 1.5 mM to obtain distance restraints (Sattler et al. 1999). For angle restraints, <sup>3</sup>J(H<sup>N</sup>H<sup>A</sup>) scalar coupling constants were determined from intensity ratios of diagonal and cross-peaks in the HNHA spectrum (Vuister and Bax 1993). Slowly exchanging amide protons were identified in a series of <sup>15</sup>N-HSQC spectra acquired after dissolving freeze-dried <sup>15</sup>N-labeled GP-NusA(339–495) in D<sub>2</sub>O, and those amide protons still visible after 15 min were assumed to be involved in hydrogen bonds. <sup>1</sup>D(NH<sup>N</sup>) RDCs were obtained from J-modulated HSQC experiments at 800 MHz (Tjandra et al. 1996) with a 0.5-mM sample of GP-NusA(339–495) in the presence of 10 mg/mL Pf1 phage (Hansen et al. 1998) by subtracting the <sup>1</sup>J(NH<sup>N</sup>) scalar coupling constants measured in reference spectra without Pf1 phage. A sample of 0.5 mM <sup>15</sup>N-labeled GP-NusA(339–495) was used for measuring <sup>15</sup>N R<sub>1</sub> and R<sub>2</sub> rates, as well as the HetNOE, with published pulse sequences (Dayie and Wagner 1994). For the determination of R<sub>1</sub> and R<sub>2</sub> at a proton frequency of 600 MHz, spectra were recorded with delays of 6.88 (3 $\times$ ), 752.33 (3 $\times$ ), 1020.48 (3 $\times$ ), and 1288.63 (3 $\times$ ) msec, and 16.96 (3 $\times$ ), 84.4 (3 $\times$ ), 101.76 (3 $\times$ ), and 118.72 (3 $\times$ ) msec, respectively.

The HetNOEs were averaged over two independent data sets. The NMR data were processed using NMRPipe (Delaglio et al. 1995) and in-house written software, and analyzed with the program package NMRview5.0.4 (Johnson and Blevins 1994). Diffusion experiments were performed at a proton resonance frequency of 400 MHz on samples of 175, 350, and 700  $\mu\text{M}$  GP-NusA(339–495) in  $\text{D}_2\text{O}$  containing 0.5% dioxan, 10 mM potassium phosphate (pH 6.8), 50 mM sodium chloride, and 0.02% sodium azide. For each experiment, five independent series of 1D spectra were acquired using the PG-SLED pulse sequence (Jones et al. 1997) with a gradient time of 7 msec, an echo time of 70 msec, and gradient strengths of 5%–50% in 5% increments, and 50%–100% in 6.25% increments. A single Gaussian was fitted to the decay of the protein signal, and the decay of the dioxane signal was described by two Gaussians to account for the overlap of the dioxane signal with the protein signal. Titrations were performed by adding aliquots of a concentrated stock solution of unlabeled  $\lambda\text{N}$ (1–53) or  $\alpha\text{CTD}$ (233–329) to a sample of  $\sim 0.5$  mM  $^{15}\text{N}$ -labeled GP-NusA(339–495) and acquiring  $^1\text{H}^{15}\text{N}$  HSQC spectra at 700 MHz after each titration step. In both titrations, ligands were added to GP-NusA(339–495) until no further changes could be observed in the spectra.

### Structure calculation

Distance restraints for structure calculation were derived from NNH-NOESY,  $^{13}\text{C}$ - and  $^{15}\text{N}$ -NOESY-HSQC spectra, and 4D-C,C-NOESY spectra. NOESY cross-peaks were classified according to their relative intensities and converted to distance restraints with upper limits of 2.7 Å (strong), 3.5 Å (medium), and 5.0 Å (weak). For ambiguous distance restraints, the  $r^{-6}$  summation over all assigned possibilities defined the upper limit (Nilges 1995).

The raw scalar coupling constants were multiplied with a correction factor of 1.1 to take into account the different relaxation rates of in-phase and anti-phase components (Vuister and Bax 1993). Residues with scalar coupling constants below 6 Hz were restrained to dihedral angles between  $-80^\circ$  and  $-40^\circ$ , residues showing coupling constants above 8 Hz were restricted to dihedral angles of  $-160^\circ$  to  $-80^\circ$  (Schweimer et al. 2002). Glycines were omitted, since they were not stereospecifically assigned, and since the coupling constants are likely affected by cross-relaxation (Vuister and Bax 1993).

Hydrogen bonds were included in the final structure calculation if the acceptor of a slowly exchanging amide proton could be identified from the results of preceding structure calculations. Thus, a hydrogen bond was assumed if the distance of the carboxyl oxygen and the amide proton was below 2.6 Å, and the angle of the amide proton, the amide nitrogen, and the carboxyl oxygen was  $< 60^\circ$  in all accepted structures. For each hydrogen bond, the distance between the amide proton and the acceptor was restrained to  $< 2.3$  Å, and the distance between the amide nitrogen and the acceptor was restrained to  $< 3.3$  Å (Schweimer et al. 2002).

All prolines were considered to adopt the *trans* conformation as strong  $\text{H}^{\text{A}}(\text{i})\text{-H}^{\text{D}}(\text{i}+1)$  and  $\text{H}^{\text{N}}(\text{i})$  and  $\text{H}^{\text{D}}(\text{i}+1)$  NOEs could be observed (Wüthrich 1986).

The structures for NusA(351–426) and NusA(426–495) were calculated separately, since no NOEs could be observed between the two regions, and relaxation data suggested an at least partly independent reorientation of the two regions. GP-NusA(339–350) was omitted in the structure calculation, due to the lack of inter-residual NOEs and low values of the HetNOEs, both indicating a flexible amino terminus.

The structure calculations were performed with the program XPLOR-NIH 1.2.1 (Schwieters et al. 2003) using a three-step simulated annealing protocol (Nilges et al. 1988a,b,c) with floating assignment of prochiral groups (Folmer et al. 1997).

First conformational space, sampling was carried out for 120 psec with a time step of 3 fsec at a temperature of 2000 K, followed by a cooling period of 120 psec down to 1000 K, and 60 psec cooling to 100 K, both with a time step of 2 fsec. A modified conformational database potential for backbone and side-chain dihedral angles was applied (Kuszewski and Clore 2000; Neudecker et al. 2001). The proline angles were modified according to Neudecker et al. (2004). After simulated annealing, the structures were subjected to 1000 steps of Powell minimization (Powell 1977), and the final 500 steps were minimized without conformational database potential.

In a first step, 90 structures were calculated (Table 1) using 883/699 [NusA(351–426)/NusA(426–495)] distance and 39/28 dihedral angle restraints; in a second step, 38/40 additional restraints from hydrogen exchange experiments were included. The 32 structures with the lowest total energy were then refined using 37/32  $^1\text{D}(\text{NH}^{\text{N}})$  RDCs with a harmonic potential (Tjandra et al. 1997). Flexible residues (for definition, see “NMR data relaxation analysis,” below) were excluded from the calculations. The tensor components of the alignment were optimized separately for NusA(351–426) and NusA(426–495) with a grid search by varying the axial component  $D_a$  and the rhombicity  $R$  in steps of 0.5 and 0.1, respectively. The initial values of  $D_a$  and  $R$  were estimated from the distribution of the  $^1\text{D}(\text{NH}^{\text{N}})$  (Clore et al. 1998) and a molecular dynamics run was performed for each pair of  $D_a$  and  $R$ , yielding axial components of  $10.5 \pm 0.25$  Hz [NusA(351–426)] and  $6.5 \pm 0.25$  Hz [NusA(426–495)], and rhombicities of  $0.2 \pm 0.05$  [NusA(351–426)] and  $0.3 \pm 0.05$  [NusA(426–495)] for the energetically most favorable combination of  $D_a$  and  $R$ .

The 19 structures showing the lowest values of the target function excluding the database potential were further analyzed with XPLOR-NIH 1.2.1 (Schwieters et al. 2003) and PROCHECK 3.5.4 (Morris et al. 1992; Laskowski et al. 1993).

Helix angles in the average structure were estimated by interpolating axes through all heavy atoms in each of the helices according to the MOLMOL least-squares algorithm.

### NMR relaxation data analysis

Relaxation rates were calculated by least-squares fitting of mono-exponential decays to the peak-heights with the program Curvefit (A.G. Palmer III, Columbia University, NY, unpubl.). Errors were estimated to be 5% for relaxation rates and 10% for the HetNOEs. The coarse filter of Pawley et al. (2001) served to identify flexible residues [six in NusA(353–416), 12 in NusA(431–490)], that is, residues with significant motions on a pico- to nanosecond timescale, and residues probably involved in exchange processes. Additionally, residues outside of regular secondary structure and H353, Y362, L373, and Y386 were not taken into account for the calculation of the diffusion tensors with Tensor2 (Dosset et al. 2000) using standard settings. Hydrodynamic parameters of an elongated NusA(353–490) molecule were computed with HYDRONMR version 5a using a shell model (de la Torre and Carrasco 2000).

### Acknowledgments

We gratefully acknowledge financial support by the Deutsche Forschungsgemeinschaft (Ro617/12-1).

## References

- Atkinson, A.R. and Kieffer, B. 2004. The role of protein motions in molecular recognition: Insights from heteronuclear NMR relaxation measurements. *Prog. Nucl. Magn. Reson. Spectrosc.* **44**: 141–187.
- Bax, A., Kontaxis, G., and Tjandra, N. 2001. Dipolar couplings in macromolecular structure determination. *Methods Enzymol.* **339**: 127–174.
- Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., and Bourne, P.E. 2000. The Protein Data Bank. *Nucleic Acids Res.* **28**: 235–242.
- Bonin, I., Muhlberger, R., Bourenkov, G.P., Huber, R., Bacher, A., Richter, G., and Wahl, M.C. 2004. Structural basis for the interaction of *Escherichia coli* NusA with protein N of phage  $\lambda$ . *Proc. Natl. Acad. Sci.* **101**: 13762–13767.
- Braddock, D.T., Cai, M., Baber, J.L., Huang, Y., and Clore, G.M. 2001. Rapid identification of medium- to large-scale interdomain motion in modular proteins using dipolar couplings. *J. Am. Chem. Soc.* **123**: 8634–8635.
- Braddock, D.T., Louis, J.M., Baber, J.L., Levens, D., and Clore, G.M. 2002. Structure and dynamics of KH domains from FBP bound to single-stranded DNA. *Nature* **415**: 1051–1056.
- Clore, G.M., Gronenborn, A.M., and Bax, A. 1998. A robust method for determining the magnitude of the fully asymmetric alignment tensor of oriented macromolecules in the absence of structural information. *J. Magn. Reson.* **133**: 216–221.
- Cordingley, M.G., Callahan, P.L., Sardana, V.V., Garsky, V.M., and Colonno, R.J. 1990. Substrate requirements of human rhinovirus 3C protease for peptide cleavage *in vitro*. *J. Biol. Chem.* **265**: 9062–9065.
- Dayie, K.T. and Wagner, G. 1994. Relaxation-rate measurements for  $^{15}\text{N}$ - $^1\text{H}$  groups with pulsed-field gradients and preservation of coherence pathways. *J. Magn. Reson.* **111A**: 121–126.
- Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J., and Bax, A. 1995. NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**: 277–293.
- de la Torre, J. and Carrasco, H.B. 2000. HYDRONMR: Prediction of NMR relaxation of globular proteins from atomic level structures and hydrodynamic calculations. *J. Magn. Reson.* **147B**: 138–146.
- Doherty, A.J., Serpell, L.C., and Ponting, C.P. 1996. The helix-hairpin-helix DNA-binding motif: A structural basis for non-sequence-specific recognition of DNA. *Nucleic Acids Res.* **24**: 2488–2497.
- Dosset, P., Hus, J.C., Blackledge, M., and Marion, D. 2000. Efficient analysis of macromolecular rotational diffusion from heteronuclear relaxation data. *J. Biomol. NMR* **16**: 23–28.
- Eisenmann, A., Schwarz, S., Rösch, P., and Schweimer, K. 2004. Sequence-specific  $^1\text{H}$ ,  $^{13}\text{C}$ ,  $^{15}\text{N}$  resonance assignments and secondary structure of the carboxy-terminal domain of the *E. coli* transcription factor NusA. *J. Biomol. NMR* **28**: 193–194.
- Folmer, R.H., Hilbers, C.W., Konings, R.N., and Nilges, M. 1997. Floating stereospecific assignment revisited: Application to an 18 kDa protein and comparison with J-coupling data. *J. Biomol. NMR* **9**: 245–258.
- Fushman, D., Varadan, R., Assfalg, M., and Walker, O. 2004. Determining domain orientation in macromolecules by using spin-relaxation and residual dipolar coupling measurements. *Prog. Nucl. Magn. Reson. Spectrosc.* **44**: 189–214.
- Gopal, B., Haire, L.F., Gamblin, S.J., Dodson, E.J., Lane, A.N., Papavinasundaram, 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.
- Gusarov, I. and Nudler, E. 2001. Control of intrinsic transcription termination by N and NusA: The basic mechanisms. *Cell* **107**: 437–449.
- Hansen, M.R., Mueller, L., and Pardi, A. 1998. Tunable alignment of macromolecules by filamentous phage yields dipolar coupling interactions. *Nat. Struct. Biol.* **5**: 1065–1074.
- Holm, L. and Sander, C. 1996. Mapping the protein universe. *Science* **273**: 595–602.
- Johnson, B.A. and Blevins, R.A. 1994. NMRview: A computer program for the visualization and analysis of NMR data. *J. Biomol. NMR* **4**: 603–614.
- Jones, J.A., Wilkins, D.K., Smith, L.J., and Dobson, C.M. 1997. Characterisation of protein unfolding by NMR diffusion measurements. *J. Biomol. NMR* **10**: 199–203.
- Kay, L.E., Torchia, D.A., and Bax, A. 1989. Backbone dynamics of proteins as studied by  $^{15}\text{N}$  inverse detected heteronuclear NMR spectroscopy: Application to staphylococcal nuclease. *Biochemistry* **28**: 8972–8979.
- Kim, C.A. and Bowie, J.U. 2003. SAM domains: Uniform structure, diversity of function. *Trends Biochem. Sci.* **28**: 625–628.
- Kohno, T., Kusunoki, H., Sato, K., and Wakamatsu, K. 1998. A new general method for the biosynthesis of stable isotope-enriched peptides using a decahistidine-tagged ubiquitin fusion system: An application to the production of mastoparan-X uniformly enriched with  $^{15}\text{N}$  and  $^{13}\text{C}$ . *J. Biomol. NMR* **12**: 109–121.
- Koradi, R., Billeter, M., and Wüthrich, K. 1996. MOLMOL: A program for display and analysis of macromolecular structure. *J. Mol. Graph.* **14**: 51–55.
- Kumar, S. and Bansal, M. 1998. Dissecting  $\alpha$ -helices: Position specific analysis of  $\alpha$ -helices in globular proteins. *Proteins* **31**: 460–476.
- Kuszewski, J. and Clore, G.M. 2000. Sources of and solution to problems in the refinement of protein NMR structures against torsion angle potentials of mean force. *J. Magn. Reson.* **146**: 249–254.
- Laskowski, R.A., MacArthur, M.W., Moss, D.S., and Thornton, J.M. 1993. PROCHECK: A program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* **26**: 283–291.
- Liu, K., Zhang, Y., Severinov, K., Das, A., and Hanna, M.M. 1996. Role of *Escherichia coli* RNA polymerase  $\alpha$  subunit in modulation of pausing, termination, and antitermination by the transcription elongation factor NusA. *EMBO J.* **15**: 150–161.
- Maciejewski, M.W., Liu, D., Prasad, R., Wilson, S.H., and Mullen, G.P. 2000. Backbone dynamics and refined solution structure of the N-terminal domain of DNA polymerase  $\beta$ . Correlation with DNA binding and dRP lyase activity. *J. Mol. Biol.* **296**: 229–253.
- 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.
- 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.
- Mandel, A.M., Akke, M., and Palmer, A.G. 1995. Backbone dynamics of *Escherichia coli* ribonuclease HI: Correlations with structure and function in an active enzyme. *J. Mol. Biol.* **246**: 144–163.
- Mogridge, J., Mah, T.F., and Greenblatt, J. 1995. A protein-RNA interaction network facilitates the template-independent cooperative assembly on RNA polymerase of a stable antitermination complex containing the  $\lambda$  N protein. *Genes & Dev.* **9**: 2831–2845.
- Mogridge, J., 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  $\lambda$  N protein. *Mol. Cell* **1**: 265–275.
- Mooney, R.A., Artsimovitch, I., and Landick, R. 1998. Information processing by RNA polymerase: Recognition of regulatory signals during RNA chain elongation. *J. Bacteriol.* **180**: 3265–3275.
- Morris, A.L., MacArthur, M.W., Hutchinson, E.G., and Thornton, J.M. 1992. Stereochemical quality of protein structure coordinates. *Proteins* **12**: 345–364.
- Neudecker, P., Sticht, H., and Rösch, P. 2001. Improving the efficiency of the Gaussian conformational database potential for the refinement of protein and nucleic acid structures. *J. Biomol. NMR* **21**: 373–375.
- Neudecker, P., Nerkamp, J., Eisenmann, A., Nourse, A., Lauber, T., Schweimer, K., Lehmann, K., Schwarzinger, S., Ferreira, F., and Rösch, P. 2004. Solution structure, dynamics, and hydrodynamics of the calcium-bound cross-reactive birch pollen allergen Bet v 4 reveal a canonical monomeric two EF-hand assembly with a regulatory function. *J. Mol. Biol.* **336**: 1141–1157.
- Nilges, M. 1995. Calculation of protein structures with ambiguous distance restraints. Automated assignment of ambiguous NOE crosspeaks and disulphide connectivities. *J. Mol. Biol.* **245**: 645–660.
- Nilges, M., Clore, G.M., and Gronenborn, A.M. 1988a. Determination of three-dimensional structures of proteins from interproton distance data by hybrid distance geometry-dynamical simulated annealing calculations. *FEBS Lett.* **229**: 317–324.
- . 1988b. Determination of three-dimensional structures of proteins from interproton distance data by dynamical simulated annealing from a random array of atoms. Circumventing problems associated with folding. *FEBS Lett.* **239**: 129–136.
- Nilges, M., Gronenborn, A.M., Brünger, A.T., and Clore, G.M. 1988c. Determination of three-dimensional structures of proteins by simulated annealing with interproton distance restraints. Application to crambin, potato carboxypeptidase inhibitor, and barley serine proteinase inhibitor 2. *Protein Eng.* **2**: 27–38.

- Notredame, C., Higgins, D., and Heringa, J. 2000. T-Coffee: A novel method for multiple sequence alignments. *J. Mol. Biol.* **302**: 205–217.
- 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 antitermination in *E. coli*. *Genes Cells* **7**: 755–768.
- Pawley, N.H., Wang, C., Koide, S., and Nicholson, L.K. 2001. An improved method for distinguishing between anisotropic tumbling and chemical exchange in analysis of  $^{15}\text{N}$  relaxation parameters. *J. Biomol. NMR.* **20**: 149–165.
- Powell, M.J.D. 1977. Restart procedures for the conjugate gradient method. *Math. Progr.* **12**: 241–254.
- Sattler, M., Schleucher, J., and Griesinger, C. 1999. Heteronuclear multi-dimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. *Prog. Nucl. Magn. Reson. Spectrosc.* **34**: 39–158.
- Schweimer, K., Hoffmann, S., Bauer, F., Friedrich, U., Kardinal, C., Feller, S.M., Biesinger, B., and Sticht, H. 2002. Structural investigation of the binding of a herpesviral protein to the SH3 domain of tyrosine kinase Lck. *Biochemistry* **41**: 5120–5130.
- Schwieters, C.D., Kuszewski, J.J., Tjandra, N., and Clore, G.M. 2003. The Xplor-NIH NMR molecular structure determination package. *J. Magn. Reson.* **160**: 66–74.
- Shao, X. and Grishin, N.V. 2000. Common fold in helix-hairpin-helix proteins. *Nucleic Acids Res.* **28**: 2643–2650.
- Shin, D.H., Nguyen, H.H., Jancarik, J., Yokota, H., Kim, R., and Kim, S.H. 2003. Crystal structure of NusA from *Thermotoga maritima* and functional implication of the N-terminal domain. *Biochemistry* **42**: 13429–13437.
- Stevens, S.Y., Sanker, S., Kent, C., and Zuiderweg, E.R. 2001. Delineation of the allosteric mechanism of a cytidyltransferase exhibiting negative cooperativity. *Nat. Struct. Biol.* **8**: 947–952.
- Thanos, C.D., Faham, S., Goodwill, K.E., Cascio, D., Phillips, M., and Bowie, J.U. 1999. Monomeric structure of the human EphB2 sterile  $\alpha$  motif domain. *J. Biol. Chem.* **274**: 37301–37306.
- Tjandra, N., Feller, S.E., Pastor, R.W., and Bax, A. 1995. Rotational diffusion anisotropy of human ubiquitin from  $^{15}\text{N}$  relaxation. *J. Am. Chem. Soc.* **117**: 12562–12566.
- Tjandra, N., Grzesiek, S., and Bax, A. 1996. Magnetic field dependence of nitrogen-proton J splittings in  $^{15}\text{N}$ -enriched human ubiquitin resulting from relaxation interference and residual dipolar coupling. *J. Am. Chem. Soc.* **118**: 6264–6272.
- Tjandra, N., Omichinski, J.G., Gronenborn, A.M., Clore, G.M., and Bax, A. 1997. Use of dipolar  $^1\text{H}$ - $^{15}\text{N}$  and  $^1\text{H}$ - $^{13}\text{C}$  couplings in the structure determination of magnetically oriented macromolecules in solution. *Nat. Struct. Biol.* **4**: 732–738.
- Tsugawa, A., Saito, M., Court, D.L., and Nakamura, Y. 1988. NusA amber mutation that causes temperature-sensitive growth of *Escherichia coli*. *J. Bacteriol.* **170**: 908–915.
- van Gilst, M.R. and von Hippel, P.H. 1997. Assembly of the N-dependent antitermination complex of phage  $\lambda$ . NusA and RNA bind independently to different unfolded domains of the N protein. *J. Mol. Biol.* **274**: 160–173.
- Vuister, W.G. and Bax, A. 1993. Quantitative J correlation: A new approach for measuring homonuclear three-bond  $J(\text{H}^{\text{NH}}^{\text{A}})$  coupling constants in  $^{15}\text{N}$ -enriched proteins. *J. Am. Chem. Soc.* **115**: 7772–7777.
- Wilkins, D.K., Grimshaw, S.B., Receveur, V., Dobson, C.M., Jones, J.A., and Smith, L.J. 1999. Hydrodynamic radii of native and denatured proteins measured by pulse field gradient NMR techniques. *Biochemistry* **38**: 16424–16431.
- 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**: 1177–11789.
- Wüthrich, K. 1986. *NMR of proteins and nucleic acids*. Wiley, New York.
- Xia, T., Frankel, A., Takahashi, T.T., Ren, J., and Roberts, R.W. 2003. Context and conformation dictate function of a transcription antitermination switch. *Nat. Struct. Biol.* **10**: 812–819.
- Zuiderweg, E.R. 2002. Mapping protein-protein interactions in solution by NMR spectroscopy. *Biochemistry* **41**: 1–7.

## **Einzelarbeit C**

Stefan Prasch, Kristian Schweimer und Paul Rösch

Structural Basis of Transcription Elongation Control: The NusA- $\alpha$ CTD Complex.

In Bearbeitung





## **Structural Basis of Transcription Elongation Control: The NusA- $\alpha$ CTD Complex**

Stefan Prasch, Kristian Schweimer, and Paul Rösch\*

Lehrstuhl für Struktur und Chemie der Biopolymere & Forschungszentrum für Bio-Makromoleküle, Universität Bayreuth, Universitätsstrasse 30, 95447 Bayreuth, Germany

\* Corresponding author; Universität Bayreuth, Lehrstuhl Biopolymere, Universitätsstr. 30, 95447 Bayreuth, Germany; Phone +49 921 55-3541; Fax: +49 921 553544; E-mail: roesch@unibt.de

Word count: 17489

Running title: NusA- $\alpha$ CTD complex

Subject Categories: structural biology; Chromatin & Transcription

Key words: transcription elongation / NusA / AR2 -  $\alpha$ CTD complex / NMR / helix-hairpin-helix

**Abstract**

Transcription in prokaryotes is mainly regulated at three distinct stages, the transcription initiation, the elongation, and the termination steps, each involving a number of proteins in addition to the RNA polymerase (RNAP) forming the active complex. Although the crystal structure of the elongating prokaryotic RNA polymerase is known, one of the crucial regulatory domains, the carboxyterminal domain of the  $\alpha$  subunit ( $\alpha$ CTD), and its interaction with other transcription factors remained elusive. The complex structure between the elongation factor NusA, acidic repeat 2, and the  $\alpha$ CTD as determined by multidimensional high resolution NMR shows a novel mode of interaction, entirely different from known structures of complexes of proteins with similar fold. NusA blocks the  $\alpha$ CTD »265 determinant« known to bind to the key upstream activating element. This suggests the NusA- $\alpha$ CTD interaction to be a key switch from initiation to elongation.

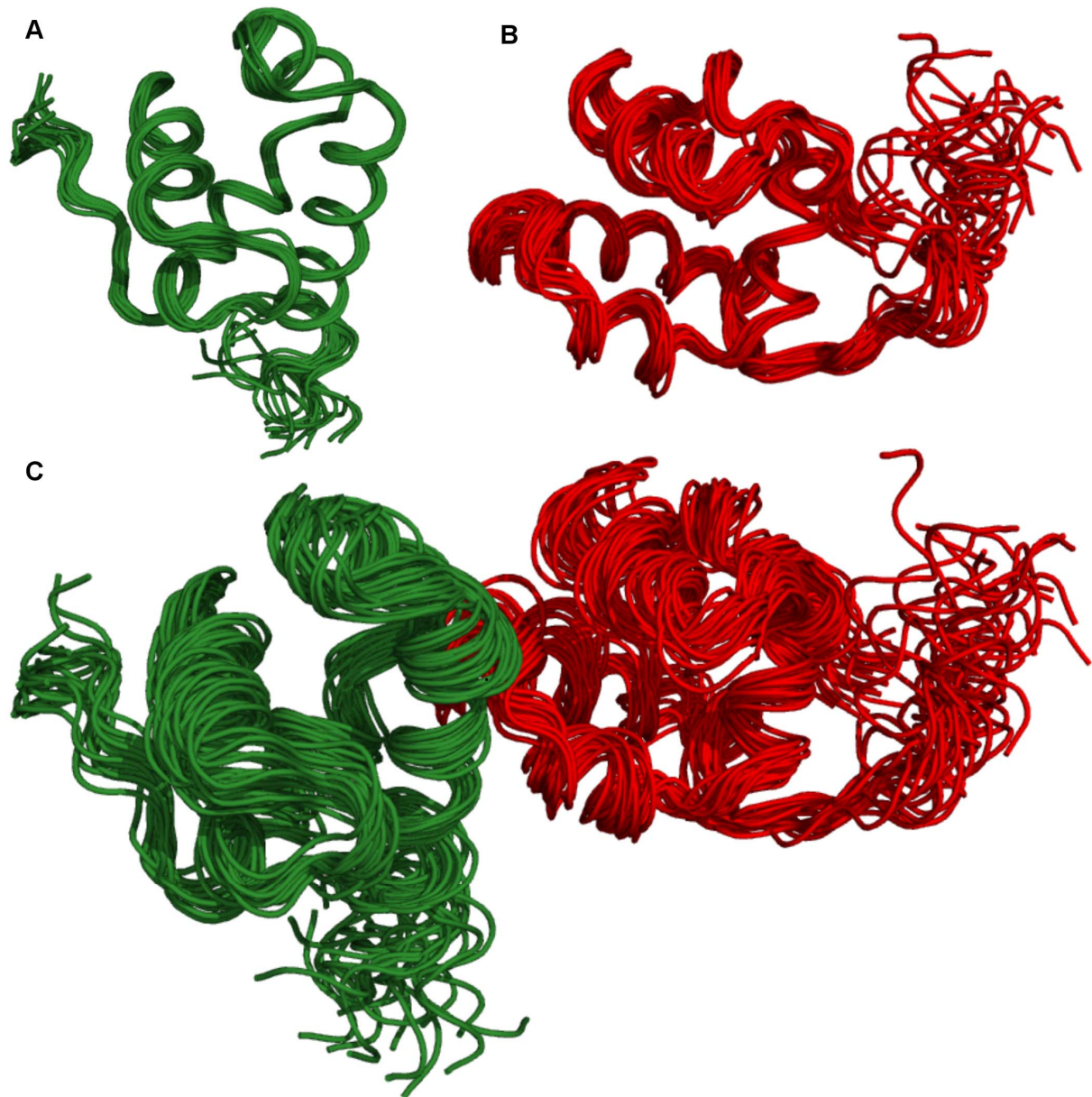
DNA-dependent RNA polymerase (RNAP) is the central enzyme of gene expression and a major target of regulatory factors. Core RNAP of *Escherichia coli* (*E. coli*) consists of five subdomains:  $\beta'$ ,  $\beta$ , a dimer of  $\alpha$  subunits and the  $\omega$  subunit<sup>1</sup>. Whereas the  $\beta$  and  $\beta'$  subunits catalyze RNA polymerization, the two  $\alpha$  subunit N-terminal domains ( $\alpha$ NTD) orchestrate RNAP assembly and the  $\alpha$  subunit C-terminus forms an independent domain ( $\alpha$ CTD) attached via an unstructured linker<sup>2</sup>. The  $\omega$  subunit is dispensable for cell viability, but may act as a chaperone and promote RNAP assembly *in vitro*<sup>3,4</sup>. The core enzyme is capable of RNA synthesis, and one of the several specificity  $\sigma$ -factors converts the core enzyme into the holoenzyme necessary to initiate transcription. Then, the  $\sigma$ -factor releases the processive RNAP<sup>5</sup>. Regulation of initiation via  $\alpha$ CTD is based on two separate events: First,  $\alpha$ CTD can interact with *cis*-acting DNA sequences upstream of promoters (upstream elements, UP), increasing the basal transcription level up to two orders of magnitude<sup>6</sup>; secondly,  $\alpha$ CTD interacts with upstream DNA bound transcription activators such as the catabolite activator protein (CAP) or the superoxide response regulator proteins MarA or SoxR<sup>7-9</sup>. Class I CAP-dependent transcription of the *lac* operon are mediated by three determinants on the surface of  $\alpha$ CTD, the »261 determinant«, residues 257-259, and 261, that mediates interaction with  $\sigma$ <sup>70</sup> factor, the »265 determinant«, residues 265, 294, 296, 299, and 302, that interacts with DNA, and the »287 determinant«, residues 285-290, 315, 317, and 318, that interacts with activators such as CAP<sup>7,10,11</sup>. In contrast, MarA and SoxS interact with  $\alpha$ CTD via the »265 Determinant«<sup>9</sup>. For successful regulation of transcription elongation several additional host proteins are required, such as N utilization substances A (NusA), B, E, and G<sup>12-15</sup>. *In vivo*, NusA is key in the regulation of transcription elongation by enhancing pausing of RNAP<sup>16-18</sup>, and NusA enhances termination of transcription at intrinsic terminators *in vitro* by interacting with nascent RNA<sup>19-21</sup>. NusA is not able to bind RNA directly as AR2 blocks the RNA binding domains of NusA<sup>22</sup>. To release this autoinhibition,  $\alpha$ CTD binds to AR2<sup>22-24</sup>.

## RESULTS

### Structure Description

We determined the solution structure of the 19 kDa complex between AR2 (residues 426-495) and  $\alpha$ CTD (residues 249-329) by heteronuclear double- and triple-resonance NMR spectroscopy. The structure calculation based on 2470 NMR-derived restraints including 33 intermolecular distance restraints resulted in ensembles with high precision of atomic coordinates for both proteins together with good stereochemical properties. No violations of distance restraints

$> 0.1 \text{ \AA}$  and dihedral angle restraints  $> 3^\circ$  could be observed. A summary of the structural and restraint statistics is given in Table 1. Figure 1 shows an overlay of the 20 lowest-energy conformers obtained from structure calculations. The superposition of  $\text{AR2}^{\text{complex}}$  in the complex shows a backbone rmsd of  $0.39 \text{ \AA}$  (Fig. 1A).  $\alpha\text{CTD}^{\text{complex}}$  alone shows a backbone rmsd of  $0.66 \text{ \AA}$  (Fig. 1B). The overlay of the complex shows a backbone rmsd of  $1.38 \text{ \AA}$  (Fig. 1C).



**FIGURE 1** Superposition of the best 20 calculated structures of AR2 in complex with  $\alpha\text{CTD}$ . Fig. 1A shows the superposition of  $\text{AR2}^{\text{complex}}$  (green) with an rmsd of  $0.39 \text{ \AA}$ . Fig. 1B shows the superposition of  $\alpha\text{CTD}^{\text{complex}}$  (red) with an rmsd of  $0.66 \text{ \AA}$ . Fig. 1C shows the superposition of the complex of AR2 (green) and  $\alpha\text{CTD}$  (red) with an rmsd of  $1.38 \text{ \AA}$ .

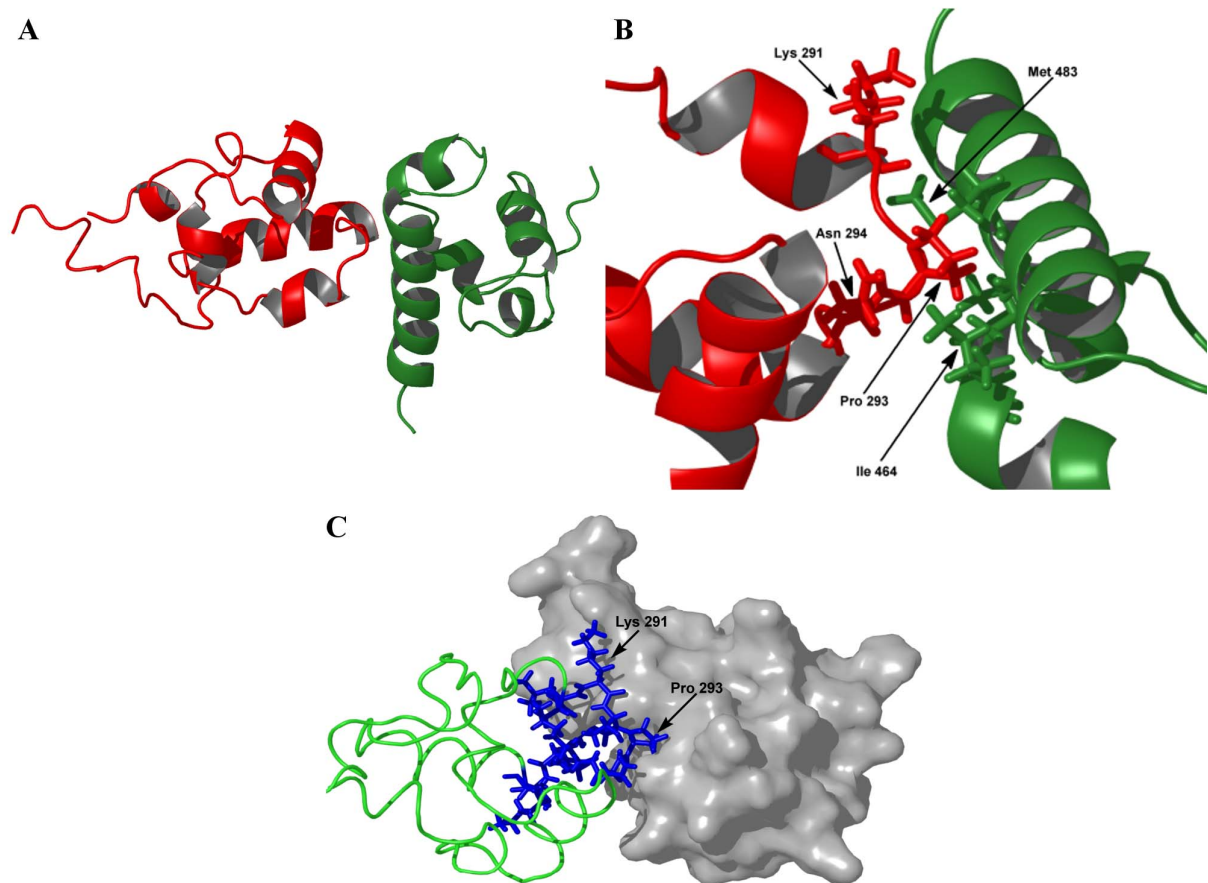
Table 1. NMR structure determination statistics.

Total NOE Distance Restraints	2470
Intermolecular (AR2 – $\alpha$ CTD)	33
Intramolecular (AR2)	1241
Intraresidual	366
Interresidual	
Sequential, medium, long	323, 330, 222
Hydrogen Bond Restraints	22
Dihedral Angle Restraints	7
Intramolecular ( $\alpha$ CTD)	1196
Intraresidual	470
Interresidual	
Sequential, medium, long	281, 244, 201
Hydrogen Bond Restraints	16
Dihedral Angle Restraints	33
<RMSD> from mean structure	
Backbone/heavy atom for AR2 – $\alpha$ CTD (Å)	1.38
Backbone/heavy atom for AR2 (Å)	0.39
Backbone/heavy atom for $\alpha$ CTD (Å)	0.66
Ramachandran Plot (% residues)	
Most Favourable Region	89.3
Additionally Allowed Region	8.5
Generously Allowed Region	2.0
Disallowed Region	0.3

Both protein subunits are predominantly  $\alpha$ -helical (Fig 2A) and each contains two helix-hairpin-helix (HhH) motifs connected by an additional  $\alpha$ -helix, an arrangement classified as a separate fold termed (HhH)<sub>2</sub><sup>25</sup>. This fold is present in a variety of protein families involved in non-sequence-specific DNA binding such as DNA-polymerases, ligases, and nucleases. In

AR2 as well as  $\alpha$ CTD, helices 1 and 2 form the first HhH-motif, and helices 4 and 5 form the second HhH-motif, helix 3 connects the individual HhH-motifs. This overall fold of AR2 and  $\alpha$ CTD does not change upon complex formation, the rmsd between free<sup>24</sup> and bound AR2 being 1.08 Å and the rmsd between free<sup>26</sup> and bound  $\alpha$ CTD being 2.00 Å for residues 249-316, nearly the entire difference arising from slightly different loop orientations (residues 293 to 296) and the slightly larger angle between helices 2 and 4 as compared to the free protein, rendering complex formation a rigid-body association mechanism<sup>27</sup>.

The interaction surface of AR2 and  $\alpha$ CTD is composed of a hydrophobic groove at the AR2 surface consisting of residues from helix 4 (Ile 464) and the centre of helix 5 (Ile 482, Met 483) and an  $\alpha$ CTD counterpart located in the hairpin (Lys291-Gly296) between helix 4 and helix 5 (Fig. 2B). Pro293 forms the tip of the  $\alpha$ CTD hairpin and is deeply imbedded into the AR2 pocket (Fig. 2C). Lys291 and probably Lys297 form salt bridges with AR2 sidechain carboxylates of AR2, but the exact counterpart could not be derived from the NMR data. The buried surface area of 811 Å<sup>2</sup> is of comparable size to that observed for other complexes involving (HhH)<sub>2</sub> domains, for example the complex between AR1 and  $\lambda$ N-peptide<sup>28</sup> and for  $\alpha$ CTD in the complex with DNA and CAP<sup>29</sup>. The small interaction interface is consistent with the finding that no large rearrangements upon complex formation take place<sup>30</sup> and relates directly to the apparent K<sub>d</sub> of AR2: $\alpha$ CTD of approximately 10  $\mu$ M as estimated from our NMR data to 100  $\mu$ M as estimated from gel shift assays<sup>22</sup>.

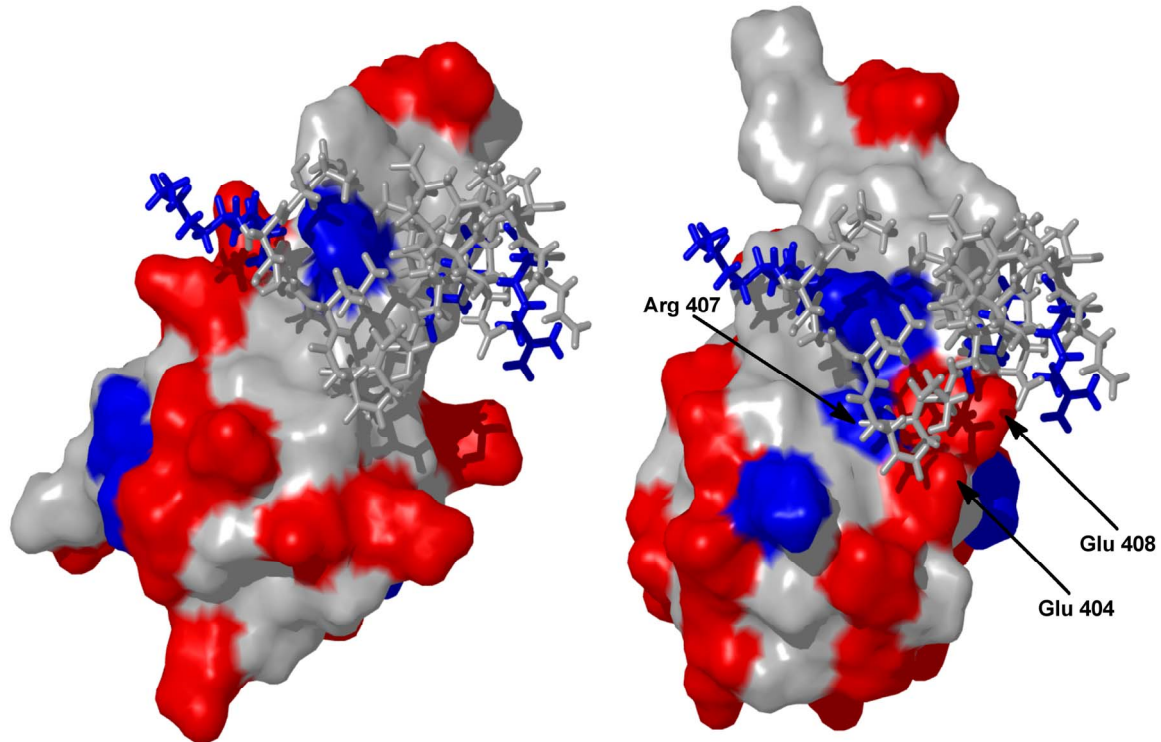


**FIGURE 2 Binding interface of AR2/αCTD complex.** Fig. 2A shows the ribbon representation of the lowest energy structure. The hairpin of αCTD (red) interacts with helix 4 and 5 of AR2 (green). Fig. 2B shows the binding interface in more detail. Pro 293 and Lys 291 of αCTD (red) interacts mainly with Met 483 and Ile 464 of AR2 (green). Fig. 2C shows the groove of AR2 (grey surface) in which Pro 293 (blue) of αCTD (green) sticks. Additional contacts arise from Lys 291 (blue).

### Structural comparison

The two carboxyterminal domains of *E. coli* NusA, AR1 and AR2, show high sequence homology, resulting in similar three dimensional structures. Nevertheless, both have different binding partners<sup>24</sup>. In the complex of AR1 with a peptide derived from λN protein AR1 displays a hydrophobic patch between helix 2 and the loop between helix 4 and helix 5 involving Phe369, Val372 and Leu398<sup>28</sup>. Leu40 from λN protein sticks into this patch forming essential contacts. Similar to the AR1: λN interaction the complex of AR2 with αCTD also displays a hydrophobic interaction surface. But here the hydrophobic patch on AR2 is located between the aminotermius of helix 4 (Ile464) and the centre of helix 5 (Ile482, Met483), located roughly at the opposite site compared to AR1: λN. The corresponding surface regions of the hydrophobic patches of one acidic repeat are distorted by a charged residue on the other acidic repeat. The presence of Lys447 in AR2 instead of the corresponding Val372 in AR1 likely

explains the specificity of  $\lambda$ N to AR1<sup>31</sup>. Projecting the binding site of AR2 for  $\alpha$ CTD onto AR1 shows that the presence of Arg407 and Glu408 in AR1 instead of Ile482 and Met483 destroys the necessary hydrophobic patch with the result of nonbinding of  $\alpha$ CTD to AR1 (Fig. 3).



**FIGURE 3 Comparison of AR1 and AR2.** Left side shows surface representation of the AR2/ $\alpha$ CTD complex. Residues 356 to 414 of AR1 have been fitted onto residue 431 to 490 of AR2 in complex with  $\alpha$ CTD (left side). Residues 280 to 300 of  $\alpha$ CTD are shown as grey sticks. Negative charged groups are coloured red, whereas positive charged groups are coloured blue. In the case of AR1 the charged residues Arg 407, Glu 404 and Glu 408 are in the centre of the binding interface interrupting the hydrophobic interaction with Pro 293. The interface of AR2 is more hydrophobic making it possible to interact with Pro 293.

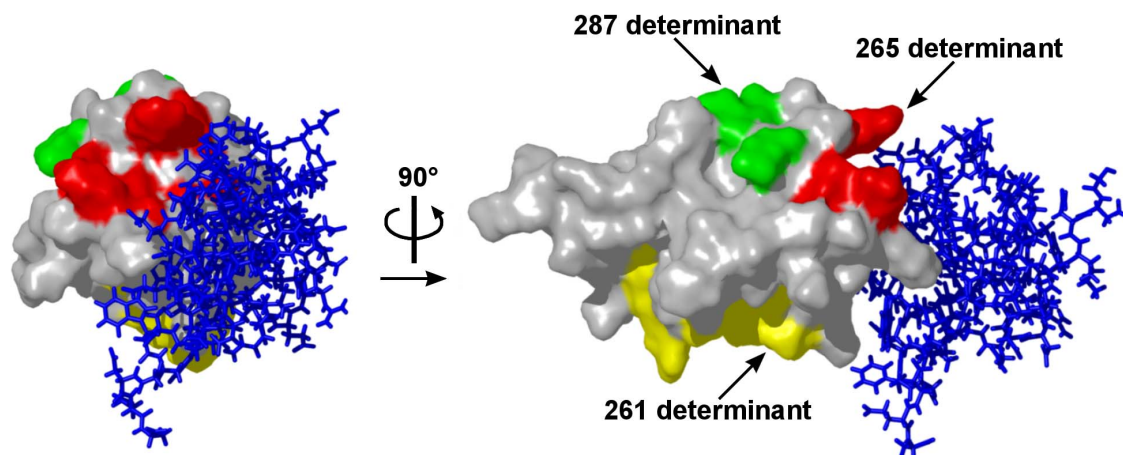
In the next step we compared the AR2: $\alpha$ CTD complex with other known complexes, consisting of two (HhH)<sub>2</sub> domains, e. g. ERCC1:XPF<sup>32,33</sup>. In contrast, the binding mode of AR2 and  $\alpha$ CTD is completely different to the ERCC1/XPF complex. In the latter case, both domains form an intermolecular four helical bundle with a pseudo-2-fold symmetry axis forming a large interaction surface. The large rearrangements reflecting a large buried surface area results in a high affinity complex. This is required to form a stable and functional heterodimeric endonuclease<sup>32,33</sup>. The interaction between NusA and the RNAP requires more transient complex formation to fulfil their regulatory function.



## DISCUSSION

The K271E in  $\alpha$ CTD impairs the efficiency of antitermination at the phage  $\lambda$   $t_R$ ,  $t_L$  and  $t_1$  terminators, presumably by decreasing the strength of the interaction between AR2 and  $\alpha$ CTD<sup>22</sup>. Indeed, K271 in helix 1 of  $\alpha$ CTD interacts with  $\alpha$ CTD Asn 268 (distance  $\sim 3\text{\AA}$ ) that in turn interacts with AR2 Asp 465 thus stabilizes the complex. A negative charge at position 271 would thus not only effect formation of helix 1, it would also interfere with the negatively charged Asp 465.

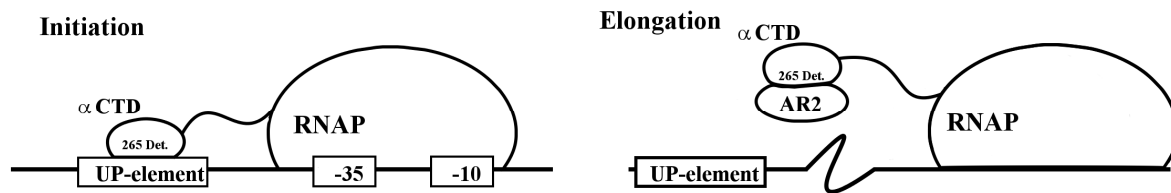
Activation of NusA is initiated by rendering its RNA binding site accessible by removal of the negatively charged AR2 that may act as an RNA mimic and bind unspecifically to the positively charged RNA binding surface of NusA. The specific hydrophobic interaction of  $\alpha$ CTD and AR2 opens the RNA binding site which, in turn, allows the S1-KH1-KH2 domains to interact with the nascent RNA and function in regulation of RNA pausing<sup>34</sup>. In addition, the interaction of  $\alpha$ CTD and the UP-element via the »265 determinant«<sup>8</sup> interferes with the binding of AR2: $\alpha$ CTD, and this determinant is also affected by the AR2 binding (Fig. 4).



**FIGURE 4 Mapping of important determinants.** Three important determinants, which have been found by biochemical and genetical studies, are mapped on the surface of  $\alpha$ CTD (grey) in complex with AR2 (blue). The determinant “261” (yellow) interacts mainly with the  $\sigma^{70}$  factor. The determinant “265” (red) interacts mainly with UP-elements, and the “287” determinant (green) interacts with Proteins like CAP. In the AR2/ $\alpha$ CTD complex the “265” determinant is blocked by AR2 (blue), whereas the “287” determinant important for protein-protein interaction is accessible.

Moreover, AR2 sterically blocks the DNA binding site of  $\alpha$ CTD. From studies with MarA or SoxS it is known that both interact with the »265 determinant«<sup>9, 35</sup>. Due to the binding of MarA or SoxS the »265 determinant« is blocked to bind to the UP-element, whereas MarA or SoxS itself could bind to the specific soxbox DNA<sup>9</sup>. Preventing  $\alpha$ CTD from binding to any transcription activation element could be seen as a switch between the transcription initiation state and the transcription elongation state (Fig. 5). The binding sites for activator proteins

(»287 determinant«) and for the  $\sigma^{70}$  factor (»261 determinant«) are accessible in the complex. The affinity of  $\sigma$  for RNAP decreases significantly after promoter escape, making  $\sigma$  release both a function of  $\sigma$ 's off rate and the effective concentration of free  $\sigma$  and of any competitors<sup>36</sup>. From the structure, both NusA and  $\sigma$  factor could be attached to  $\alpha$ CTD simultaneously and the »287 determinant« could act as a recognition site for elongation factors even in the presence of NusA. Thus, from the present results, the role of NusA in transcription regulation could be even more central than hitherto accepted.



**FIGURE 5 Schematic representations of the proposed effect of AR2 on control of transcription.** RNAP binds to the promoter indicated by the -35 and -10 elements. During initiation  $\alpha$ CTD binds via the »265 determinant« to the UP-element. During elongation AR2 binds to  $\alpha$ CTD and releases latter from the UP-element.

## Materials and Methods

### Sample preparation

*E. coli* AR1 (residues 339-426) and AR2 (residues 424-495) were expressed as a deca-histidine-tagged protein in BL21(DE3) from pET-19b (Novagen) and purified as described previously<sup>24, 31</sup>. *E. coli*  $\alpha$ CTD (residues 233-329) was expressed as a deca-histidine-tagged protein in BL21(DE3) from pET-19b and purified as described previously<sup>24</sup>. Briefly, all used constructs were grown in M9 minimal medium supplemented with  $^{15}\text{NH}_4\text{Cl}$  and 0.2%  $^{13}\text{C}$  D-glucose as the sole nitrogen and carbon sources, respectively<sup>37</sup>. The recombinant protein was purified under native conditions by nickel-affinity chromatography (5 mL His-trap chelating column, GE Health Care) and eluted by applying an imidazole step gradient. Peak fractions containing the constructs were cleaved off using PreScission protease and dialyzed against NMR buffer. NMR samples contained 0.5-1.5 mM AR1, AR2, or  $\alpha$ CTD in 10 mM potassium phosphate (pH 6.8), 50 mM sodium chloride, 0.02% sodium azide, and 10%  $\text{D}_2\text{O}$ . For complex sample we used a stoichiometric ratio of AR2: $\alpha$ CTD of 1:3, or vice versa.

### NMR measurements

All NMR experiments were conducted at 298 K on Bruker Avance 700 MHz or Bruker Avance 800 MHz spectrometers equipped with a cryogenic probe. Standard double- and triple resonance through-bond correlation experiments were recorded for resonance assignment<sup>38</sup>.

$^{13}\text{C}$ - and  $^{15}\text{N}$  spin echo difference experiments were performed for  $\chi_1$ -restraints<sup>39</sup>.  $^{15}\text{N}$ - and  $^{13}\text{C}$ -edited NOESY experiments (mixing time 120 ms) were acquired for obtaining distance restraints.  $^{13}\text{C}$  isotope-filtered NOESY experiments were recorded for both samples to identify intermolecular NOE cross peaks<sup>40</sup>. NMR data were processed using in-house written software and visualized with NMRView<sup>41</sup>.

### Structure calculation

Distance restraints for structure calculation were derived from 3D isotope edited or filtered NOESY spectra. NOESY cross peaks were classified according to their relative intensities and converted to distance restraints with upper limits of 3.0 Å (strong), 4.0 Å (medium), 5.0 Å (weak), and 6.0 Å (very weak). For ambiguous distance restraints the  $r^{-6}$  summation over all assigned possibilities defined the upper distance limit.

Hydrogen bonds were included for backbone amide protons in regular secondary structure if the amide proton did not show a water exchange cross peak in the  $^{15}\text{N}$ -edited NOESY spectrum. Dihedral restraints for  $\chi_1$ -angles based on  $^3\text{J}(\text{N},\text{H}^\beta)$ ,  $^3\text{J}(\text{N},\text{CH}_3)$ , and  $^3\text{J}(\text{CO},\text{CH}_3)$  coupling constants were included as restraints for the most probable rotamer with 20° tolerance. The structure calculations were performed with the program XPLOR-NIH 1.2.1 using a three-step simulated annealing protocol with floating assignment of prochiral groups including a conformational database potential<sup>42</sup>.

The 20 structures showing the lowest values of the target function excluding the database potential were further analyzed with X-PLOR<sup>42</sup>, MolMol<sup>43</sup>, and PROCHECK 3.5.4<sup>44</sup>. All figures were made using PYMOL<sup>45</sup>. The structure coordinates were deposited in the protein data bank (pdb, accession code XXX), and chemical shift assignments were deposited in the BioMagResBank (BMRB, accession code XXX).

### Acknowledgment

This project was supported by a grant to PR from the Deutsche Forschungsgemeinschaft DFG (Ro617/16-1).

## References

1. Severinov, K. RNA polymerase structure-function: insights into points of transcriptional regulation. *Curr. Opin. Microbiol.* **3**, 118-125 (2000).
2. Jeon, Y. H., Yamazaki, T., Otomo, T., Ishihama, A. & Kyogoku, Y. Flexible linker in the RNA polymerase alpha subunit facilitates the independent motion of the C-terminal activator contact domain. *J. Mol. Biol.* **267**, 953-962 (1997).
3. Mukherjee, K., Nagai, H., Shimamoto, N. & Chatterji, D. GroEL is involved in activation of Escherichia coli RNA polymerase devoid of the omega subunit in vivo. *Eur. J. Biochem.* **266**, 228-235 (1999).
4. Mukherjee, K. & Chatterji, D. Studies on the omega subunit of Escherichia coli RNA polymerase--its role in the recovery of denatured enzyme activity. *Eur. J. Biochem.* **247**, 884-889 (1997).
5. Nechaev, S. & Severinov, K. Bacteriophage-induced modifications of host RNA polymerase. *Annu. Rev. Microbiol.* **57**, 301-322 (2003).
6. Ross, W. *et al.* A third recognition element in bacterial promoters: DNA binding by the alpha subunit of RNA polymerase. *Science* **262**, 1407-1413 (1993).
7. Zou, C., Fujita, N., Igarashi, K. & Ishihama, A. Mapping the cAMP receptor protein contact site on the alpha subunit of Escherichia coli RNA polymerase. *Mol. Microbiol.* **6**, 2599-2605 (1992).
8. Benoff, B. *et al.* Structural basis of transcription activation: the CAP-alpha CTD-DNA complex. *Science* **297**, 1562-1566 (2002).
9. Shah, I. M. & Wolf, R. E., Jr. Novel protein--protein interaction between Escherichia coli SoxS and the DNA binding determinant of the RNA polymerase alpha subunit: SoxS functions as a co-sigma factor and redeploys RNA polymerase from UP-element-containing promoters to SoxS-dependent promoters during oxidative stress. *J. Mol. Biol.* **343**, 513-532 (2004).
10. Murakami, K., Fujita, N. & Ishihama, A. Transcription factor recognition surface on the RNA polymerase alpha subunit is involved in contact with the DNA enhancer element. *EMBO J.* **15**, 4358-4367 (1996).
11. Savery, N. J. *et al.* Transcription activation at Class II CRP-dependent promoters: identification of determinants in the C-terminal domain of the RNA polymerase alpha subunit. *EMBO J.* **17**, 3439-3447 (1998).
12. Ward, D. F. & Gottesman, M. E. The nus mutations affect transcription termination in Escherichia coli. *Nature* **292**, 212-215 (1981).
13. Mason, S. W., Li, J. & Greenblatt, J. Host factor requirements for processive antitermination of transcription and suppression of pausing by the N protein of bacteriophage lambda. *J. Biol. Chem.* **267**, 19418-19426 (1992).
14. Liu, K. & Hanna, M. M. NusA contacts nascent RNA in Escherichia coli transcription complexes. *J. Mol. Biol.* **247**, 547-558 (1995).

15. Squires, C. L., Greenblatt, J., Li, J., Condon, C. & Squires, C. L. Ribosomal RNA anti-termination in vitro: requirement for Nus factors and one or more unidentified cellular components. *Proc. Natl. Acad. Sci. U. S. A.* **90**, 970-974 (1993).
16. Lau, L. F., Roberts, J. W. & Wu, R. RNA polymerase pausing and transcript release at the lambda tR1 terminator in vitro. *J. Biol. Chem.* **258**, 9391-9397 (1983).
17. Morgan, E. A. Antitermination mechanisms in rRNA operons of Escherichia coli. *J. Bacteriol.* **168**, 1-5 (1986).
18. Landick, R. The regulatory roles and mechanism of transcriptional pausing. *Biochem. Soc. Trans.* **34**, 1062-1066 (2006).
19. Greenblatt, J., McLimont, M. & Hanly, S. Termination of transcription by nusA gene protein of Escherichia coli. *Nature* **292**, 215-220 (1981).
20. Liu, K. & Hanna, M. M. NusA interferes with interactions between the nascent RNA and the C-terminal domain of the alpha subunit of RNA polymerase in Escherichia coli transcription complexes. *Proc. Natl. Acad. Sci. U. S. A.* **92**, 5012-5016 (1995).
21. Liu, K. & Hanna, M. M. NusA contacts nascent RNA in Escherichia coli transcription complexes. *J. Mol. Biol.* **247**, 547-558 (1995).
22. Mah, T. F., Kuznedelov, K., Mushegian, A., Severinov, K. & Greenblatt, J. The alpha subunit of E. coli RNA polymerase activates RNA binding by NusA. *Genes Dev.* **14**, 2664-2675 (2000).
23. Liu, K., Zhang, Y., Severinov, K., Das, A. & Hanna, M. M. Role of Escherichia coli RNA polymerase alpha subunit in modulation of pausing, termination and anti-termination by the transcription elongation factor NusA. *EMBO J.* **15**, 150-161 (1996).
24. Eisenmann, A., Schwarz, S., Prash, S., Schweimer, K. & Rosch, P. The E. coli NusA carboxy-terminal domains are structurally similar and show specific RNAP- and lambdaN interaction. *Protein Sci.* **14**, 2018-2029 (2005).
25. Shao, X. & Grishin, N. V. Common fold in helix-hairpin-helix proteins. *Nucleic Acids Res.* **28**, 2643-2650 (2000).
26. Jeon, Y. H. *et al.* Solution structure of the activator contact domain of the RNA polymerase alpha subunit. *Science* **270**, 1495-1497 (1995).
27. Wodak, S. J. & Janin, J. Structural basis of macromolecular recognition. *Adv. Protein Chem.* **61**, 9-73 (2002).
28. Bonin, I. *et al.* Structural basis for the interaction of Escherichia coli NusA with protein N of phage lambda. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 13762-13767 (2004).
29. Benoff, B. *et al.* Structural basis of transcription activation: the CAP-alpha CTD-DNA complex. *Science* **297**, 1562-1566 (2002).
30. Lo Conte, L., Chothia, C. & Janin, J. The atomic structure of protein-protein recognition sites. *J. Mol. Biol.* **285**, 2177-2198 (1999).

31. Prasad, S. *et al.* Interaction of the intrinsically unstructured phage lambda N Protein with Escherichia coli NusA. *Biochemistry* **45**, 4542-4549 (2006).
32. Tsodikov, O. V., Enzlin, J. H., Scharer, O. D. & Ellenberger, T. Crystal structure and DNA binding functions of ERCC1, a subunit of the DNA structure-specific endonuclease XPF-ERCC1. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 11236-11241 (2005).
33. Tripsianes, K. *et al.* The structure of the human ERCC1/XPF interaction domains reveals a complementary role for the two proteins in nucleotide excision repair. *Structure* **13**, 1849-1858 (2005).
34. Nudler, E. & Gusarov, I. Analysis of the intrinsic transcription termination mechanism and its control. *Methods Enzymol.* **371**, 369-382 (2003).
35. Dangi, B., Gronenborn, A. M., Rosner, J. L. & Martin, R. G. Versatility of the carboxy-terminal domain of the alpha subunit of RNA polymerase in transcriptional activation: use of the DNA contact site as a protein contact site for MarA. *Mol. Microbiol.* **54**, 45-59 (2004).
36. Mooney, R. A., Darst, S. A. & Landick, R. Sigma and RNA polymerase: an on-again, off-again relationship? *Mol. Cell* **20**, 335-345 (2005).
37. Marley, J., Lu, M. & Bracken, C. A method for efficient isotopic labeling of recombinant proteins. *J. Biomol. NMR* **20**, 71-75 (2001).
38. Sattler, M., Schleucher, J. & Griesinger, C. Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. *Prog. Nucl. Magn. Reson. Spectrosc.* **34**, 39-158 (1999).
39. Bax, A. *et al.* Measurement of homo- and heteronuclear J couplings from quantitative J correlation. *Methods Enzymol.* **239**, 79-105 (1994).
40. Zwahlen, C. *et al.* Methods for Measurement of Intermolecular NOEs by Multinuclear NMR Spectroscopy: Application to a Bacteriophage  $\phi$  N-Peptide/boxB RNA Complex. *J. Am. Chem. Soc.* **119**, 6711-6721 (1997).
41. Johnson, B. A. & Blevins, R. A. NMRview: A computer program for the visualization and analysis of NMR data. *J. Biomol. NMR* **4**, 603-614 (1994).
42. Tjandra, N., Omichinski, J. G., Gronenborn, A. M., Clore, G. M. & Bax, A. Use of dipolar  $^1\text{H}$ - $^{15}\text{N}$  and  $^1\text{H}$ - $^{13}\text{C}$  couplings in the structure determination of magnetically oriented macromolecules in solution. *Nat. Struct. Biol.* **4**, 732-738 (1997).
43. Koradi, R., Billeter, M. & Wuthrich, K. MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graph.* **14**, 51-5, 29-32 (1996).
44. Laskowski, R. A., Rullmann, J. A., MacArthur, M. W., Kaptein, R. & Thornton, J. M. AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR* **8**, 477-486 (1996).
45. DeLano, W. L. The PyMOL Molecular Graphics System. *DeLano Scientific, Palo Alto, CA, USA.* (2002).

## Einzelarbeit D

Stefan Prasch, Marcel Jurk, Robert S. Washburn, Max E. Gottesman, Birgitta M. Wöhrl und Paul Rösch

*E. coli* NusA recognizes phage  $\lambda$  *nut spacer* RNA sequences.

Eingereicht in Journal of Biological Chemistry (2007)





## *E. coli* NusA recognizes phage $\lambda$ *nut spacer* RNA sequences

Stefan Prasch, Marcel Jurk, Robert S. Washburn§, Max E. Gottesman§, Birgitta M. Wöhrl, and Paul Rösch\*

Lehrstuhl für Biopolymere & Research Center for Bio-Macromolecules, Universität Bayreuth, Universitätsstraße 30, 95447 Bayreuth, Germany

§Department of Microbiology and Institute of Cancer Research, Columbia University Medical Center, New York, New York 10032

Running title: NusA antitermination

\* Corresponding author; Universität Bayreuth, Lehrstuhl Biopolymere, Universitätsstr. 30, 95447 Bayreuth, Germany; Phone +49 921 55-3541; Fax: +49 921 553544; E-mail: roesch@unibt.de

The RNA sequences *boxA*, *boxB*, and *boxC* nucleate the formation of the termination-resistant RNA polymerase complex on the  $\lambda$  chromosome. The complex includes *E. coli* proteins NusA, NusB, NusG, and NusE, and the  $\lambda$  N protein. A similar complex forms at the *rrn* leader. Whereas RNA-binding by NusB and NusE has been described in quantitative terms, the interaction of NusA with these RNA-sequences is less defined. From isotropic and anisotropic fluorescence equilibrium titration, we show that the main binding site for NusA is the spacer between *nut boxA* and *nut boxB*. Thus, *nutR boxA5-spacer*, *nutR boxA16-spacer*, and *nutR boxA69-spacer* retain NusA binding, whereas a mutation in the *spacer* eliminates complex formation between *nutL-boxA-spacer* and NusA. Furthermore, the *nutR spacer* alone binds NusA. NusA binds with different affinities to the  $\lambda$  *nutL*, and  $\lambda$  *nutR* sequences. The NusA:  $\lambda$  *nutL boxA-spacer* complex  $K_d$ -value is 50 % that of the corresponding  $\lambda$  *nutR* complex. In contrast, *rrn boxA*, which includes an additional U residue, binds NusA in the absence of *spacer*. The  $K_d$ -values obtained for *rrn boxA* and *rrn boxA-spacer* are 19-fold and 8-fold lower, respectively, than those for *nutR boxA-spacer*.

Gene expression in *E. coli* and its phage can be controlled at the level of transcription termination. The best-studied exam-

ples of this mechanism are the *rrn* operons and the bacteriophage  $\lambda$ .  $\lambda$  and other lambdoid phages express N, an RNA binding protein of the arginine rich motif (ARM) family, which modifies RNA polymerase (RNAP) so that transcription termination is suppressed specifically on  $\lambda$  templates. This modification, known as »antitermination«, occurs at two sites on the phage chromosome, *nutL* and *nutR*, where a complex is formed that includes N, RNAP, and four *E. coli* factors, NusA, NusB, NusE, and NusG (1; see 2). The *nut* sites consist of 4 elements, *boxA*, *spacer*, *boxB*, and *boxC*, transcribed in that order. *nut* RNA is known to bind NusA, NusB, and NusE. *boxB* RNA forms a stem-loop that binds N and NusA. *boxC* consists of an alternating GT sequence located downstream of *boxB* in lambdoid *nut* sites (3). The roles of *spacer* RNA and *boxC* have not yet been demonstrated. However, two short oligo-ribonucleotides derived from the *boxC* stem-loop motif bind exclusively to the two KH domains of NusA in an entirely extended conformation. Adenine-backbone interactions with the trinucleotide sequence AUA are particularly critical (4). The *nut* sequences and the Nus factors are also utilized by the phage HK022 Nun protein to arrest transcription on the  $\lambda$  chromosome (5-7). Transcription of the *E. coli rrn* operons is controlled by antitermination (8). The leader sequences of *rrn* operons carry a stem-loop element, followed by *boxA* and *boxC*. A second *boxA* is located in the spacer region between the

16S and 23S genes. *rrn boxA* is sufficient to suppress transcription termination in *rrn* operons. The sequence of the *rrn boxA* (5'-**ugcucuua**-3') differs from that of  $\lambda$  *boxA* (5'-**cgucuua**-3') and is thought to be more efficient in promoting antitermination. The same Nus factors that mediate  $\lambda$  N antitermination are operative in the *rrn* operons, in addition to S4 and other host factors (8). As in the  $\lambda$  antitermination system, the Nus factors form a ribonucleoprotein complex which bestows termination resistance upon RNAP in the *rrn* operons. In the phage  $\lambda$  *nut* system, however, antitermination requires the  $\lambda$  N protein as an additional component (9). This may correspond to the fact that *rrn* antitermination is effective only at rho-dependent terminators, whereas  $\lambda$  antitermination complexes are highly resistant to both rho-dependent and rho-independent terminators (10). In both cases, modified RNAP retains the ability to transcribe through multiple terminators.

NusA is an essential bacterial transcription factor in wild-type *E. coli* (11, 12). In addition to promoting antitermination, it is involved in RNA polymerase (RNAP) pausing (13, 14) and transcription termination (15, 16). These reactions may be promoted by contacts between NusA and the 3'OH end of nascent RNA (17). NusA consists of five functional subdomains: an N-terminal domain that interacts with RNAP (18), three RNA binding domains, S1, KH1, and KH2 (4, 19, 20), and two C-terminal acidic domains, AR1 and AR2, specific to the *E. coli* protein that interact with antitermination factor N and RNAP (21). AR2 masks the three RNA binding domains, thereby preventing NusA interaction with RNA (22). Structures of homologous NusA proteins from *Thermotoga maritima* (*T. maritima*) and from *Mycobacterium tuberculosis* (*M. tuberculosis*) were determined in the absence and presence of RNA, respectively (4, 19, 20). Both structures show NusA to be highly elongated (10, 23). Although knowledge of NusA function increased in recent years, several key ques-

tions concerning NusA's molecular interactions are still open: Does *E. coli* NusA bind specifically or non-specifically to RNA? Is the *boxA* sequence or the *spacer* RNA between *boxA* and *boxB* at the *nut* site key, or are both equally important for NusA binding? Is there any structural difference between the complex of NusA with *rrn* RNA and the complex of NusA formed with *nut* RNA?

## Materials and Methods

*Buffers and reagents* – All fluorescence titrations were performed in 50 mM potassium phosphate, 100 mM NaCl, 10 mM  $\beta$ -mercaptoethanol, pH 7.6, unless otherwise stated. Oligodeoxynucleotides as well as fluorescence labeled oligoribonucleotides were obtained from biomers.net (Ulm, Germany; Tab. 1) and used according to the manufacturer's instructions.

*Plasmid construct, expression, and protein purification* – A synthetic gene adapted to the *E. coli* codon usage harboring the DNA sequence of the NusA RNA binding domains from amino acid 132 to 348 was cloned *via* the BamHI and NdeI restriction sites into the *E. coli* expression vector pET11a (Novagen). The soluble recombinant protein contained an N-terminal 5xHis tag. NusA was expressed and purified according to published procedures (22). Briefly, *E. coli* strain BL21 (DE3) (Novagen) harboring the recombinant plasmid was grown at 37 °C in LB medium (Luria-Bertani) containing ampicillin (100  $\mu$ g/ml) until an OD<sub>600</sub> = 0.5 was reached and then induced by 0.1 mM isopropyl 1-thio- $\beta$ -D-galactopyranoside (IPTG). Cells were harvested 4 h after induction, lysed, and purified as described (22). Finally, the protein was dialyzed against buffer as used for fluorescence measurements (50 mM sodium-phosphate, 100 mM NaCl, 10 mM  $\beta$ -mercaptoethanol, pH 7.6). The dialyzed protein was concentrated with Vivaspin concentrators (Vivascience, MWCO 10,000 Da). The identity and structural integrity of purified protein was analyzed

by 19 % SDS- PAGE as well as by CD- and NMR spectroscopy.

*NMR spectroscopy* - NMR spectra were recorded on Bruker DRX 600 MHz spectrometers with triple-resonance probes equipped with pulsed field-gradient capabilities. The sample temperature was 298 K. 1D <sup>1</sup>H spectra were collected with water suppression using a 1-1 spin echo pulse sequence including gradients.

*Fluorescence equilibrium measurements* – We used various RNA sequences corresponding to the *nut* region of the  $\lambda$  genome or to the *rrnG boxA* sequence (*rrn BoxA*) of the *E. coli* genome (Table 1). Fluorescence equilibrium titrations were performed using an L-format Jobin-Yvon Horiba Fluoromax fluorimeter equipped with an automatic titration device (Hamilton). Extrinsic fluorescence measurements with 3' 6-carboxy-fluorescein (6-FAM) labeled RNA were performed in fluorescence buffer as above in a total volume of 1 ml using a 10 x 4 mm quartz cuvette (Hellmanex). The excitation wavelength was 492 nm, and the emission intensity was measured at 516 nm applying a 500 nm cutoff filter. For isotropic measurements slit widths were set at 4 nm and 2 nm for excitation and emission, respectively; for anisotropic measurements slit widths were set at 4.5 nm and 3.5 nm for excitation and emission, respectively. All titration measurements were performed at 25 °C with 50 nM of fluorescence labeled RNA. Following sample equilibration, at least 6 data points with an integration time of 0.8 s were collected for each titration point in the case of anisotropic measurements.

*Data fitting* – Isotropic as well as anisotropic data were fitted to a two-state bind-

ing equation to determine the equilibrium dissociation constant ( $K_d$ ) using standard software. The anisotropy was calculated from:

$$A = f_{\text{complex}} A_{\text{complex}} + f_{\text{RNA}} A_{\text{RNA}} \quad (1)$$

where  $A$ ,  $A_{\text{complex}}$ , and  $A_{\text{RNA}}$  are the anisotropy values and  $f_{\text{complex}}$ ,  $f_{\text{RNA}}$  are the fractional intensities. The change in fluorescence intensity has to be taken into account, so that the bound fraction is given by

$$\frac{[\text{complex}]}{[\text{RNA}]_0} = \frac{A - A_{\text{RNA}}}{A - A_{\text{RNA}} + R(A_{\text{complex}} - A)} \quad (2)$$

with

$$[\text{complex}] = \frac{K_D + [P]_0 + [\text{RNA}]_0 - \sqrt{(K_D + [P]_0 + [\text{RNA}]_0)^2 - 4[P]_0[\text{RNA}]_0}}{2[\text{RNA}]_0} \quad (3)$$

where  $A$  is the anisotropy;  $A_{\text{RNA}}$  is the initial free anisotropy,  $A_{\text{complex}}$  is the anisotropy of the protein-RNA complex and  $P_0$  and  $\text{RNA}_0$  represent the total protein and RNA concentrations, respectively.  $R$  is the ratio of intensities of the bound and free forms.

For isotropic measurements Equation 4 was used,

$$F = F_{\text{max}} - (F_{\text{max}} - F_{\text{min}}) \cdot \frac{\sqrt{(K_D + [P]_0 + [\text{RNA}]_0)^2 - 4[P]_0[\text{RNA}]_0}}{2[\text{RNA}]_0} \quad (4)$$

where  $F$  is the fluorescence signal;  $F_{\text{max}}$  and  $F_{\text{min}}$  represent the signals for the unbound and bound form,  $P_0$  and  $\text{RNA}_0$  are the total protein and RNA concentrations, respectively.

Table 1

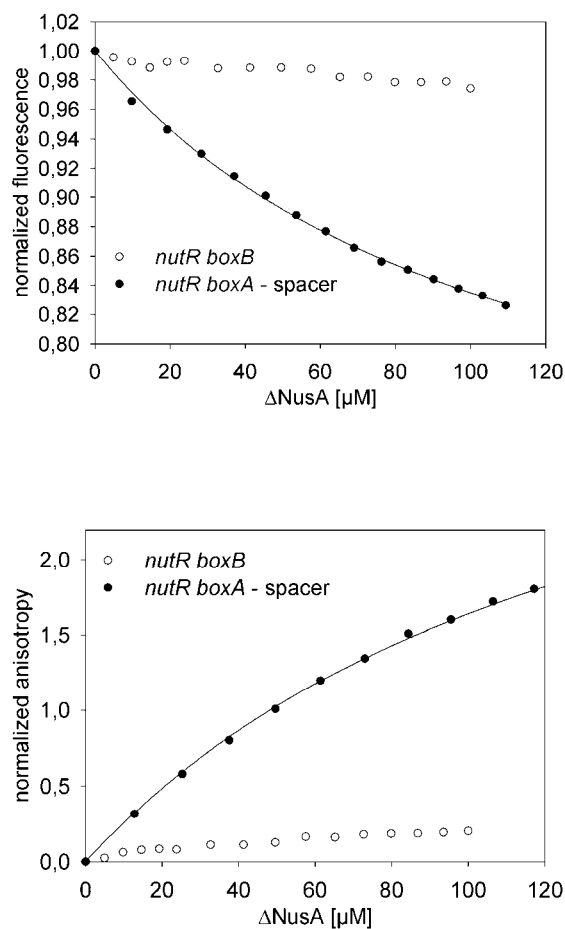
Oligonucleotide	Sequence	$K_D$ for NusA complex [ $\mu\text{M}$ ]
<i>nutR boxB</i>	5'-agcccugaaaaagggc-3'	n.d.
<i>nutR boxA-spacer</i>	5'- <b>cg</b> <u>cucuu</u> acacaucca-3'	126 $\pm$ 4
<i>nutR boxA</i>	5'- <b>cg</b> <u>cucuu</u> a-3'	n.d.
<i>nutR spacer</i>	5'-cacaucca-3'	137 $\pm$ 17
<i>nutR CAC region</i>	5'- <b>uu</b> acacauu-3'	n.d.
<i>nutR boxA5-spacer</i>	5'- <b>cu</b> <u>cucuu</u> acacaucca-3'	124 $\pm$ 7
<i>nutR boxA16-spacer</i>	5'- <b>cg</b> <u>cu</u> <u>uu</u> acacaucca-3'	106 $\pm$ 4
<i>nutR boxA69-spacer</i>	5'- <b>au</b> <u>agcggc</u> cacaucca-3'	n.d.
<i>nutL boxA-spacer</i>	5'- <b>cg</b> <u>cucuu</u> aaaauuu-3'	71 $\pm$ 4
<i>nutL boxA-spacer (mut)</i>	5'- <b>cg</b> <u>cucuu</u> aaaaaggaa-3'	n.d.
<i>rrn boxA-spacer</i>	5'-cac <u>ugcucuu</u> uuaacaauua-3'	14 $\pm$ 0.16

## Results

To determine the interaction of *E. coli* NusA with different RNA substrates, we used a NusA construct lacking the N-terminal domain and the two acidic-repeat domains ( $\Delta\text{NusA}$ ), since these regions are not directly involved in RNA binding. AR2 masks the RNA binding domains of NusA (18, 22). Withdrawal of AR2 by binding to RNAP  $\lambda$  subunit, or by deletion of the domain, allows NusA to bind RNA (22, 24). Indeed, *E. coli* strains expressing a truncated NusA lacking the AR1/AR2 domains are viable (25).

*$\Delta\text{NusA}$  binds to the nut site spacer*– The *nutR* recognition site is composed of three distinct regions, *boxA*, *boxB*, and a connecting sequence of nine bases (*spacer*). (Table 1) Upon addition of  $\Delta\text{NusA}$ , the isotropic fluorescence signal of *nutR boxA-spacer* significantly decreased in intensity (Fig. 1A). The two component binding isotherm fitting procedure yielded a  $K_d$  of  $\sim 109 \mu\text{M}$ . Titration of *nutR boxB* with  $\Delta\text{NusA}$ , in contrast, showed no, or only very weak, non-specific protein-RNA interactions (Fig 1A).

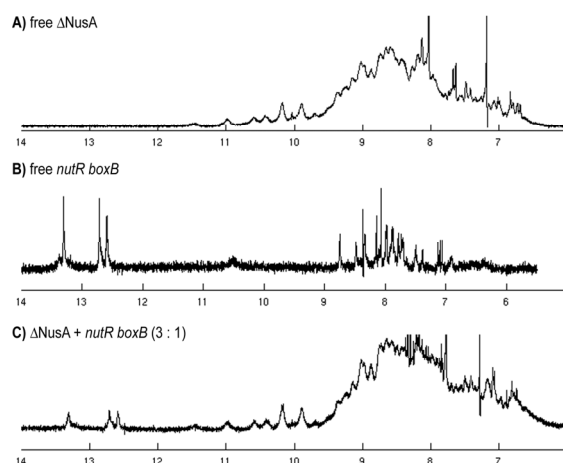
One disadvantage of isotropic fluorescence measurements is the possibility of a false negative result due to binding by  $\Delta\text{NusA}$  too distal to the fluorescence dye to alter the fluorescence signal intensity. To avoid such an artifact, we performed anisotropic fluorescence titrations. This method has proven useful for detecting molecular interactions, even if the signal change in isotropic fluorescence is weak or absent (26). Further, with anisotropic measurements changes of the fluorophore environment can be neglected since the results are related to the rotational correlation time of a macromolecule with a rigidly attached fluorophore (27). Evaluation of the results of the anisotropic fluorescence measurement yielded a virtually identical  $K_d$ -value of  $\sim 126 \mu\text{M}$  for the titration of *nutR boxA-spacer* with  $\Delta\text{NusA}$ . As before, no change of fluorescence signal intensity was observed for *nutR boxB* titration with  $\Delta\text{NusA}$  (Fig.1B; Table 1).



**FIGURE 6 Fluorescence equilibrium titrations with  $\Delta$ NusA.** 50 nM *nutR boxB* (open circles) and *nutR boxA-spacer* (filled circles) was titrated with  $\Delta$ NusA. The extrinsic fluorescence of the 3' 6-FAM label of the RNAs was determined. **(A)** Isotropic fluorescence. The lines show the best fit to eq. 2 (Materials and Methods). A  $K_d$ -value of 109  $\mu$ M was determined for *nutR boxA-spacer* (solid line). **(B)** Anisotropic fluorescence. The curves show the best fit to equation 3 (Materials & Methods). A  $K_d$ -value of 126  $\mu$ M was determined for *nutR boxA* plus spacer (solid line).

To confirm that *nutR boxB* does not interact with  $\Delta$ NusA even at higher concentrations, we analyzed a sample containing both compounds with 1D-NMR. In contrast to *nutR boxA*, *nutR boxB* forms a stable stem-loop structure allowing the detection of the slowly exchanging imino protons in the double-stranded stem region. The 1D-NMR spectrum of  $\Delta$ NusA in the absence of RNA shows a well-dispersed

amide proton signal region, indicating a stably folded, highly structured protein (Fig. 2A). The *nutR boxB* 1D-NMR spectrum reveals signals in the range of 12 to 14 ppm, corresponding to the imino protons of the stem region (Fig. 2B). Interaction between the stem region of *nutR boxB* and  $\Delta$ NusA, would affect these readily observable imino proton signals. The observable signals, however, of both, protein and RNA, are unchanged in a sample containing  $\Delta$ NusA and *nutR boxB*, clearly indicating that no complex forms between  $\Delta$ NusA and *nutR boxB* even at NusA concentrations in the high micromolar range. The observed signal increase is due to lower concentration of *nutR* RNA upon addition of  $\Delta$ NusA.



**FIGURE 7 1D-NMR analysis.** Imino proton region of  $\Delta$ NusA (175  $\mu$ M; A), *nutR boxB* (100  $\mu$ M; B), and  $\Delta$ NusA + *nutR boxB* (3:1; C).

*The nutR spacer is important for specific binding of  $\Delta$ NusA* – After our initial result that  $\Delta$ NusA binds to *nutR boxA-spacer*, but poorly or not at all to *nutR boxB* we proceeded to further define the NusA interaction region at *nutR*. We examined three smaller segments of the *nutR boxA-spacer* sequence (Table 1): *nutR boxA*, the *spacer*, and the *nutR boxA-spacer* junction, i.e. the terminal 3 nucleotides of *nutR boxA* and 6 nt of the *spacer* (the CAC region). Fluorescence anisotropy measurements revealed that *nutR boxA-spacer* binds

$\Delta$ NusA with the highest affinity ( $K_d \sim 126 \mu\text{M}$ ; Fig. 3), whereas  $\Delta$ NusA binding to *nutR boxA* could not be detected. Interestingly, the *nutR spacer* binds  $\Delta$ NusA with an affinity nearly identical to that of *nutR boxA-spacer* ( $K_d$ -value  $\sim 137 \mu\text{M}$ ). The CAC region showed only very weak  $\Delta$ NusA interaction, and no reliable  $K_d$ -value could be determined. These results clearly show that the main contribution to  $\Delta$ NusA binding at *nutR* is from the *spacer* (Fig. 3).

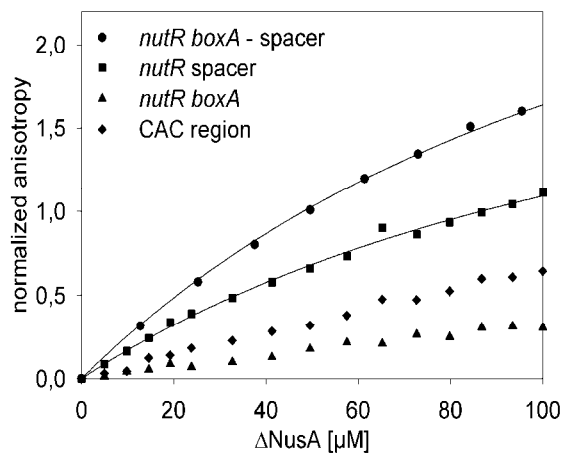


FIGURE 8 **Fluorescence anisotropy measurements with different *nut* RNAs.** 50 nM of 6-FAM labeled *nutR boxA-spacer* (circles), *nutR spacer* (squares), *nutR boxA* (triangles), or the CAC-region (diamonds) was titrated with  $\Delta$ NusA.  $K_d$ -values of 126  $\mu\text{M}$  and 137  $\mu\text{M}$  were determined for *nutR boxA-spacer* and *nutR spacer*, respectively (solid lines). No  $K_d$ -values could be fitted to *nutR boxA* or the CAC-region (see table).

To validate this result, we analyzed *nutR boxA-spacer* sequences with mutations in the *boxA* region, known to affect Nun-mediated termination and  $\lambda$  antitermination (6, 28). The *boxA5* and *boxA16* mutations decrease N activity, whereas the *boxA69* mutation has little effect on antitermination (29). Fluorescence titrations of the three mutant RNAs indicate that only *boxA69* significantly increased the  $K_d$ -value ( $>200 \mu\text{M}$ ) for  $\Delta$ NusA complex formation, whereas *boxA5* and *boxA16* exhibited  $K_d$ -values similar to that of wild type *boxA* ( $\sim 120 \mu\text{M}$ ; Fig. 4). This data suggests that

mutations inside the *boxA* sequence that affect antitermination have only a very limited effect on  $\Delta$ NusA binding. Their phenotype instead may reflect a failure to bind NusB (30). Why *nutR boxA69-spacer* binds NusA less efficiently than *nutR spacer* alone is unclear.

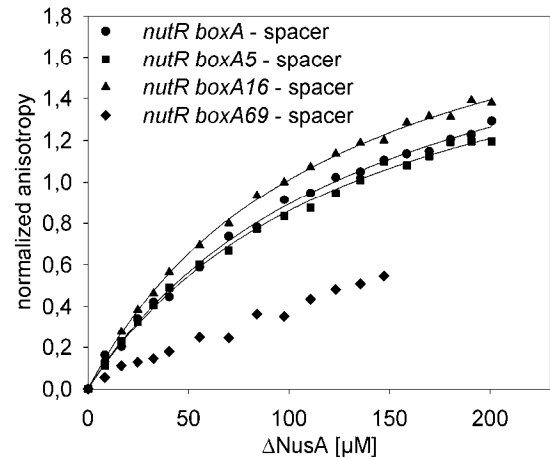


FIGURE 9 **Fluorescence anisotropy measurements with mutated *nutR* BoxA RNAs.** 50 nM of 6-FAM labeled *nutR boxA-spacer* (circles), *nutR boxA5-spacer* (squares), *nutR boxA16-spacer* (triangles), or *nutR boxA69-spacer* (diamonds) was titrated with  $\Delta$ NusA.  $K_d$ -values of 126  $\mu\text{M}$ , 124  $\mu\text{M}$ , 106  $\mu\text{M}$  were determined, respectively (solid lines; see table). No  $K_d$ -values could be fitted to *nutR boxA69* (see table).

The  $\lambda$  *nutL spacer* sequence differs from that of  $\lambda$  *nutR spacer* (Table 1), and this difference is thought to account at least partially for the enhanced efficiency of Nun-mediated termination at *nutL* relative to *nutR* (R.S.W. and M.E.G., unpublished), and *nutL boxA-spacer* binds with significantly higher affinity to  $\Delta$ NusA ( $K_d = 71 \mu\text{M}$ ) than does *nutR boxA-spacer* ( $K_d = 126 \mu\text{M}$ ). Both *spacer* sequences contain uracils at residues 5 and 6, suggesting these bases to be important for binding of interaction partners. Indeed, transversion of these residues to guanine completely abolished NusA binding to *nutL-spacer*.

*rrn boxA-spacer* exhibits the highest affinity to  $\Delta$ NusA – We next turned our atten-

tion to the *rrn* operons which are also regulated at the level of antitermination by NusA and the other Nus factors (8). *rrn* carries *boxA* and *boxC* sequences. The *boxA* sequence of *rrn* differs from that of  $\lambda$  at the initial base and by the insertion of an additional U residue at the penultimate site, converting the *rrn boxA* to a consensus site. Conversion of  $\lambda boxA$  to consensus enhances N activity (28). The *spacer* sequence of *rrn* differs from both *nutL* and *nutR*, but all three spacers carry nucleotides U6 and U7. We find that the *rrn boxA-spacer* binds with higher affinity to  $\Delta$ NusA ( $K_d = 14 \mu\text{M}$ ; Fig. 5; Table 1) than either the  $\lambda nutR$ -*spacer* or the  $\lambda nutL$ -*spacer*.

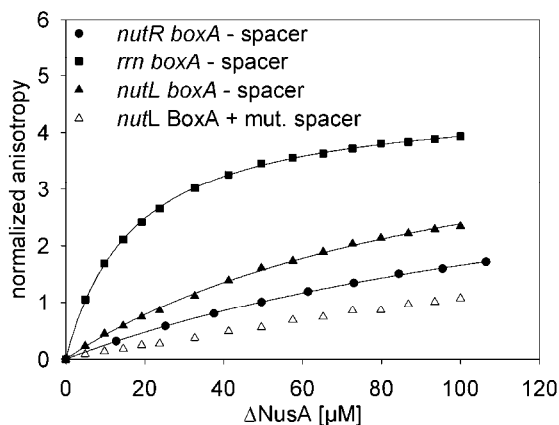


FIGURE 10 **Fluorescence anisotropy measurements with homologous *nut*-RNAs.** 50 nM of 6-FAM labeled *nutR boxA* plus spacer (filled circles), *rrn boxA* (squares), *nutL boxA* plus spacer (filled triangles), or *nutL boxA* plus mutated spacer (open triangles) were titrated with  $\Delta$ NusA.  $K_d$ -values are shown summarized in Table2. No  $K_d$ -value could be determined for *nutL boxA* plus mutated spacer.

**Conclusions** - Mutational studies indicated that NusA as well as *nutR boxA* play an important role in antitermination (31, 32). *boxA* forms a complex with NusB/NusE (30, 33) and it was suggested that NusA links *nut boxA* and *nut boxB* by binding to both (30, 34). Oddly, however, and in contrast to *boxA* point mutations, antitermination was still efficient - and NusB-independent - in a *boxA* deletion mu-

tant(29, 35). Additionally, deletion of the initial three bases (**cac**) of the *spacer* did not affect antitermination, whereas deletion of the initial six bases (**cacauu**) led to complete loss of antitermination activity. NusA interaction with the *rrn*, *nutL* and *nutR boxA-spacer* motif was demonstrated by mutational studies and *in vitro* binding assays (3, 30, 32, 36), and NusA was also suggested to recognize RNA outside the *nut*-region (12, 19, 30, 37). Complex formation between NusA and *spacer* would easily allow for binding of NusE/NusB to the adjacent *boxA* sequence, and the notion that NusA binds to the *nut spacer* region is now strongly supported by the present fluorescence titration data.

Differences between *rrn* and *nutR* have already been described (30, 33). Unlike *rrn boxA* which is sufficient to prevent termination at Rho-dependent terminators, antitermination at *nutR* requires  $\lambda$  N and *boxB* (30, 38). While the NusE/NusB complex binds with affinities in the nano-molar range (33), the  $K_d$ -values obtained for  $\Delta$ NusA are in the micro-molar range, which might reflect that, in contrast to NusB, NusA is already bound to the RNAP and thus is necessarily in close vicinity to the nascent RNA.

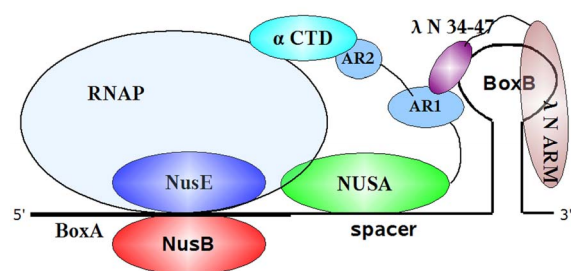


FIGURE 11 **Model of the antitermination network.** The interaction of the RNAP with various factors important for antitermination (see text for details.)

From this and other data, the following picture of the assembly of the antitermination complex at the *nut* RNA evolves (Fig. 6): After RNAP has synthesized the *nut* site, NusE and NusB bind to *boxA*, and NusA binds to the *spacer*. In the presence

of phage  $\lambda$  the  $\lambda$ N protein is expressed and N binds to AR1 of NusA as demonstrated for N(34-47), forming a weak helix at the protein's N-terminus (39). This weak helix facilitates recognition of *boxB*, and NusA AR2 interacts with the C-terminal domain of the  $\alpha$  subunit of the RNAP. NusA interaction with RNA is thus stabilized by the AR1:RNAP interaction as well as by the AR2:N:*boxB* interaction, making a very tight binding of NusA to *nut* unnecessary.

### REFERENCES

- Friedman, D. I., and Baron, L. S. (1974) *Virology* **58**, 141-148
- Nudler, E., and Gottesman, M. E. (2002) *Genes Cells* **7**, 755-768
- Berg, K. L., Squires, C., and Squires, C. L. (1989) *J. Mol. Biol.* **209**, 345-358
- Beuth, B., Pennell, S., Arnvig, K. B., Martin, S. R., and Taylor, I. A. (2005) *EMBO J.* **24**, 3576-3587
- Robert, J., Sloan, S. B., Weisberg, R. A., Gottesman, M. E., Robledo, R., and Harbrecht, D. (1987) *Cell* **51**, 483-492
- Robledo, R., Gottesman, M. E., and Weisberg, R. A. (1990) *J. Mol. Biol.* **212**, 635-643
- Robledo, R., Atkinson, B. L., and Gottesman, M. E. (1991) *J. Mol. Biol.* **220**, 613-619
- Squires, C. L., Greenblatt, J., Li, J., Condon, C., and Squires, C. L. (1993) *Proc. Natl. Acad. Sci. U. S. A.* **90**, 970-974
- Das, A., and Wolska, K. (1984) *Cell* **38**, 165-173
- Condon, C., Squires, C., and Squires, C. L. (1995) *Microbiol. Rev.* **59**, 623-645
- Zhou, Y., Mah, T. F., Yu, Y. T., Mogridge, J., Olson, E. R., Greenblatt, J., and Friedman, D. I. (2001) *J. Mol. Biol.* **310**, 33-49
- Zhou, Y., Mah, T. F., Greenblatt, J., and Friedman, D. I. (2002) *J. Mol. Biol.* **318**, 1175-1188
- Landick, R., and Yanofsky, C. (1987) *J. Mol. Biol.* **196**, 363-377
- Chan, C. L., and Landick, R. (1993) *J. Mol. Biol.* **233**, 25-42
- Farnham, P. J., Greenblatt, J., and Platt, T. (1982) *Cell* **29**, 945-951
- Schmidt, M. C., and Chamberlin, M. J. (1987) *J. Mol. Biol.* **195**, 809-818
- Liu, K., and Hanna, M. M. (1995) *J. Mol. Biol.* **247**, 547-558
- Mah, T. F., Li, J., Davidson, A. R., and Greenblatt, J. (1999) *Mol. Microbiol.* **34**, 523-537
- Worbs, M., Bourenkov, G. P., Bartunik, H. D., Huber, R., and Wahl, M. C. (2001) *Mol. Cell* **7**, 1177-1189
- Gopal, B., Haire, L. F., Gamblin, S. J., Dodson, E. J., Lane, A. N., Papavinasasundaram, K. G., Colston, M. J., and Dodson, G. (2001) *J. Mol. Biol.* **314**, 1087-1095
- Eisenmann, A., Schwarz, S., Prasch, S., Schweimer, K., and Rosch, P. (2005) *Protein Sci.* **14**, 2018-2029
- Mah, T. F., Kuznedelov, K., Mushegian, A., Severinov, K., and Greenblatt, J. (2000) *Genes Dev.* **14**, 2664-2675
- Weisberg, R. A., and Gottesman, M. E. (1999) *J. Bacteriol.* **181**, 359-367
- Arnvig, K. B., Pennell, S., Gopal, B., and Colston, M. J. (2004) *Proc. Natl. Acad. Sci. U. S. A.* **101**, 8325-8330
- Bylund, G. O., Lovgren, J. M., and Wikstrom, P. M. (2001) *J. Bacteriol.* **183**, 6095-6106
- Takahashi, M., Sakumi, K., and Sekiguchi, M. (1990) *Biochemistry* **29**, 3431-3436
- Cantor, C. R., and Schimmel, P. R. (1981) **Part II**, pp 454-465
- Olson, E. R., Tomich, C. S., and Friedman, D. I. (1984) *J. Mol. Biol.* **180**, 1053-1063
- Zuber, M., Patterson, T. A., and Court, D. L. (1987) *Proc. Natl. Acad. Sci. U. S. A.* **84**, 4514-4518
- Mogridge, J., Mah, T. F., and Greenblatt, J. (1995) *Genes Dev.* **9**, 2831-2845
- Olson, E. R., Flamm, E. L., and Friedman, D. I. (1982) *Cell* **31**, 61-70
- Friedman, D. I., and Olson, E. R. (1983) *Cell* **34**, 143-149



33. Greive, S. J., Lins, A. F., and von Hippel, P. H. (2005) *J. Biol. Chem.* **280**, 36397-36408
34. Mogridge, J., Legault, P., Li, J., Van Oene, M. D., Kay, L. E., and Greenblatt, J. (1998) *Mol. Cell* **1**, 265-275
35. Patterson, T. A., Zhang, Z., Baker, T., Johnson, L. L., Friedman, D. I., and Court, D. L. (1994) *J. Mol. Biol.* **236**, 217-228
36. Schauer, A. T., Carver, D. L., Bigelow, B., Baron, L. S., and Friedman, D. I. (1987) *J. Mol. Biol.* **194**, 679-690
37. Tsugawa, A., Kurihara, T., Zuber, M., Court, D. L., and Nakamura, Y. (1985) *EMBO J.* **4**, 2337-2342
38. Morgan, E. A. (1986) *J. Bacteriol.* **168**, 1-5
39. Prasch, S., Schwarz, S., Eisenmann, A., Wohrl, B. M., Schweimer, K., and Rosch, P. (2006) *Biochemistry* **45**, 4542-4549

### FOOTNOTES

\* This project was supported by a grant to BMW and PR from the Deutsche Forschungsgemeinschaft DFG (Ro617/16-1) and to MEG from the NIH (NIGMS R01/GM37219). PR would like to thank the Columbia University Microbiology group for their patience during his sabbatical stay in New York.

<sup>1</sup>Abbreviations are: RNAP, RNA polymerase; Nus, N-utilization substance; nut, N-utilization site; *rrn* ribosomal RNA;



## Danksagung

Die vorliegende Arbeit wurde von September 2004 bis Dezember 2007 am Lehrstuhl für Struktur und Chemie der Biopolymere der Universität Bayreuth unter der Leitung von Prof. Dr. Paul Rösch angefertigt. Ihm möchte ich für die Überlassung des sehr interessanten Themas und der hervorragenden Arbeitsmöglichkeiten sowie sein Interesse am Fortgang und Gelingen der Arbeit danken.

Mein ganz besonderer Dank gilt Dr. Kristian Schweimer und Prof. Dr. Birgitta Wöhrl für ihre Unterstützung und Diskussionsbereitschaft bei der vorliegenden Arbeit.

Bedanken möchte ich mich auch bei Hanna Berkner, Björn Burmann, Anke Eisenmann, Andrea Hager, Maximilian Hartl, Ramona Heissmann, Nadine Herz, Marcel Jurk, Christian Mangels, Irena Matecko, Ulrike Persau, Sujatha Pagadala Santhanam, Stephan Schwarzinger und Sabine Wenzel für die ausgesprochen gute Laboratmosphäre. Gudrun Wagner und Violaine Zigan danke ich für die Unterstützung in den Verwaltungsangelegenheiten. Ebenso möchte ich Rainer Hofmann für die vielen Tipps im Umgang mit dem Computer danken. Insbesondere möchte ich Anke Eisenmann danken, die mich bestens und mit viel Hingabe in das interessante Forschungsgebiet eingearbeitet hat.

Herzlichen Dank an Prof. Dr. Max E. Gottesman (University of Columbia, USA) für den sehr abwechslungsreichen und interessanten Aufenthalt in seiner Arbeitsgruppe. Ebenso möchte ich Dr. Robert S. Washburn danken, der mich in die Geheimnisse der Bakteriengenetik einführte.



## **Erklärung**

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbständig verfasst und keine 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