

**Die Regulation der Transkriptionsantitermination in
*Escherichia Coli***

Dissertation

zur Erlangung des Doktorgrades

der Fakultät für Biologie, Chemie und Geowissenschaften

der Universität Bayreuth

Vorgelegt von

Diplom Biochemiker

Björn Marcus Burmann

aus Rheda-Wiedenbrück

Bayreuth 2009

Die vorliegende Arbeit wurde von September 2006 bis Juli 2009 am Lehrstuhl Biopolymere und Forschungszentrum für Bio-Makromoleküle unter der Leitung von Prof. Dr. Paul Rösch erstellt.

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

Promotionsgesuch eingereicht am: 28.07.2009

Tag des wissenschaftlichen Kolloquiums: 18.03.2010

Prüfungsausschuss:

Prof. Dr. Paul Rösch (Erstgutachter)

Prof. Dr. Matthias Ullmann (Zweitgutachter)

Prof. Dr. Andreas Fery (Vorsitzender)

Prof. Dr. Carlo Unverzagt

Inhaltsverzeichnis

Inhaltsverzeichnis.....	I
Kurzfassung	II
Summary.....	IV
1 Einleitung.....	1
1.1 Der Phage λ	1
1.2 Lytischer und lysogener Zyklus	1
1.3 Bakterielle Transkription	3
1.3.1 Die bakterielle RNA-Polymerase	3
1.3.2 Initiation	8
1.3.3 Elongation.....	10
1.3.4 Termination	12
1.4 Antitermination.....	15
1.4.1 Das Antiterminationssystem des Phagen λ	16
1.4.2 Die ribosomale (<i>rrn</i>) Antitermination.....	19
1.5 Die Termination des Phagen HK022.....	20
1.6 Die Nus-Elongationsfaktoren.....	21
1.6.1 NusA	22
1.6.2 NusB und NusE (S10).....	23
1.6.3 NusG	25
2 Ziele.....	28
3 Zusammenfassung und Diskussion der Ergebnisse.....	30
3.1 Strukturelle Untersuchung von molekularen Komplexen.....	30
3.2 Die Bindung an die Schleife der <i>boxB</i> ist für Nun nicht essentiell	31
3.3 Die Interaktion des NusB:NusE Heterodimers mit RNA	33
3.4 Die einzelnen NusG-Domänen zeigen transiente Interaktionen	35
3.5 NusG rekrutiert das NusB:NusE Heterodimer zur RNAP	36
3.6 Implikationen für die Transkriptions-Translations-Kopplung	38
4 Abkürzungsverzeichnis.....	40
5 Literaturverzeichnis.....	42
6 Publikationsliste	58
6.1 Einzelarbeit A.....	58
6.2 Einzelarbeit B.....	58
6.3 Einzelarbeit C.....	59
6.4 Einzelarbeit D.....	59
6.5 Einzelarbeit E.....	60
6.6 Einzelarbeit F	60
Einzelarbeit A	61
Einzelarbeit B.....	69
Einzelarbeit C.....	76
Einzelarbeit D	85
Einzelarbeit E.....	100
Einzelarbeit F.....	115
Danksagung	151
Erklärung	153

Kurzfassung

Die Transkription wird in den Zellen durch die RNA-Polymerasen (RNAP) katalysiert. Die bakterielle RNAP ist aus den α_2 , β' , β , ω -Untereinheiten aufgebaut. Das katalytische Zentrum wird durch die β' - und β -Untereinheit gebildet und sorgt für die Bildung der Phosphodiesterbindung. Die α - und ω -Untereinheiten sind für den korrekten Aufbau der RNAP verantwortlich, wobei Teile der α -Untereinheiten zusätzlich regulatorische Funktionen haben. Der Ablauf der Transkription ist in Initiation, Elongation und Termination unterteilt. Alle drei Phasen werden innerhalb der Zellen streng reguliert. Während der Elongation bewirkt die sogenannte Antitermination das Überlesen von Terminationssequenzen. Dieser zuerst beim Phagen λ identifizierte Mechanismus, beruht auf dem Aufbau eines Multi-Protein-Komplexes, bestehend aus den *Escherichia coli* (*E. coli*) Nus-Proteinen (A, B, E, G), der RNAP, der *nut* RNA-Sequenz und dem λ N-Protein. Die ribosomalen (*rrn*) Operons in *E. coli* werden durch intrinsische Antitermination reguliert.

Fluoreszenz-Anisotropie-Messungen konnten zeigen, dass die Bildung des NusB:NusE Heterodimers und seine Bindung an die *boxA* der *nut* und *rrn* RNA-Sequenzen ein wichtiger Schritt im Aufbau des Antiterminationskomplexes ist. Für die sogenannte NusB101-Mutation, Asp118Asn, die den negativen Effekt von Mutationen in anderen Komponenten des Komplexes aufheben kann, konnte eine höhere Affinität zur Erkennungs-RNA gezeigt werden. Weitere Mutationen an dieser Position zeigten, dass die Aminosäureposition 118 zentral für die Stabilität der RNA-Bindung ist.

Ein weiterer Bestandteil des Antiterminationskomplex ist das zwei Domänenprotein NusG, über dessen Interaktion innerhalb des Komplexes wenig bekannt war. Interaktionen für die beiden Domänen, die das NusG-Paralog RfaH konstituieren, sind in der Literatur beschrieben, wie es auch bei der Kristallisation von NusG aus *Aquifex aeolicus* gezeigt werden konnte. Für *E. coli* NusG konnte in dieser Arbeit gezeigt werden, dass diese Interaktion konzentrationsabhängig, damit intermolekular, und sehr kurzlebig ist. Zusätzlich konnten Proteinbindungsflächen auf beiden Domänen bestimmt werden.

Als Erkenntnis von zentraler Wichtigkeit und sehr weitreichenden Konsequenzen wurde in dieser Arbeit eine spezifische und strukturell sehr gut definierte Wechselwirkung der carboxy-terminalen Domäne von NusG mit NusE, das unter der Bezeichnung S10 auch ein Bestandteil der 30S Untereinheit des Ribosoms ist, identifiziert. Seit wenigen Wochen ist in der Literatur die Interaktion der aminoterminalen Domäne von NusG mit der RNA-Polymerase beschrieben. Die hier beobachtete Wechselwirkung der carboxy-terminalen Domäne von NusG mit NusE könnte damit zur Rekrutierung des NusB:NusE Heterodimers und zur Ausbildung eines kompakten Antiterminationskomplexes beitragen. Da NusE auch Teil des Ribosoms ist, stellt die NusG:NusE Interaktion möglicherweise die erste direkte molekulare Verbindung zwischen Transkription und Translation in Bakterien dar. Diese Kopplung von bakterieller Transkription und Translation ist ein sehr lange bekanntes, aber molekular nicht erklärtes Phänomen von außerordentlicher Bedeutung für die Überlebensfähigkeit von Bakterien.

In ähnlicher Weise, wie das λ N-Protein, aber mit entgegengesetzter Wirkung – Termination statt Antitermination – bindet das Nun-Protein aus dem Phagen HongKong 022 an die *boxB*, eine Haarnadelschleife der *nut* Sequenz. Durch *in vivo* Studien wurde die Nun Tyr39Ala Mutante, die wichtige Hinweise für das Verständnis der Peptid-RNA-Interaktion lieferte, identifiziert. Durch diese Mutation ist eine ursprünglich als wichtig angesehene Wechselwirkung mit dem Adenosin 9 der *boxB* nicht mehr möglich. Molekulardynamikuntersuchungen zeigen deutliche Unterschiede zur λ N:*boxB* Wechselwirkung auf, da die analoge Trp18:A9 Interaktion essentiell für prozessive Antitermination ist. Auf diese Weise lieferte diese Untersuchung wichtige Hinweise auf den molekularen Schaltmechanismus.

Summary

The RNA-Polymerase (RNAP), the multidomain complex that catalyzes transcription, consists of the α_2 , β' , β , ω -subunits, where the β' - and β -subunits form the active center that catalyzes the RNA polymerization. The α - and ω -subunits account for the correct folding and assembly of the RNAP, whereas parts of the α -subunits perform additional regulatory functions. Transcription is subdivided into initiation, elongation, and termination, and all three phases are under tight control within the cells. Antitermination, a mechanism initially identified in the phage λ /*Escherichia coli* (*E. coli*) guest/host system, enables RNAP to read through transcription termination sites during elongation. Antitermination depends on the assembly of a multi protein complex, consisting of the *E. coli* Nus-factors A, B, E, and G, the RNAP, the *nut* RNA sequence, and the λ N-Protein. Ribosomal (*rrn*) operons in *E. coli* are also regulated by antitermination.

It could be shown by fluorescence anisotropy measurements that the formation of the NusB:NusE heterodimer and its binding to *boxA* of the *nut* and *rrn* RNA-sequences is a crucial step in the assembly of the antitermination complex. For the so-called NusB101 mutant, Asp118Asn, which is a gain of function mutant that is able to compensate for nonfunctional mutations in other Nus-factors, tighter binding than wt to RNA could be demonstrated. Results obtained with additional mutants underscored that position 118 is crucial for the stability of the RNA binding.

Another part of the antitermination complex is the two-domain protein NusG. About NusG's role within the complex little is known so far. Interactions between the two domains that constitute the NusG paralog RfaH are described in the literature, and a similar observation was described in *Aquifex aeolicus* NusG crystals. It could be shown here that for *E. coli* NusG this interaction is concentration dependent, that it is intermolecular, and only transiently populated. In addition protein-binding sites on both domains could be described.

Central point to this work and with far-reaching consequences is the specific and structurally well-defined interaction between the carboxy-terminal domain of NusG and NusE. NusE is also part of the 30S subunit of the ribosome as ribosomal protein

S10. The interaction between the amino-terminal domain of NusG and the RNAP was recently described in the literature. The observed interaction between the carboxy-terminal domain of NusG and NusE could recruit the NusB:NusE heterodimer and thus account for the assembly of a compact antitermination complex. As NusE is also part of the ribosome, this NusG:NusE interaction represents the first direct molecular link between transcription and translation in bacteria. This coupling of bacterial transcription and translation is a very well known phenomenon that, however, is not well understood, although it is of extraordinary importance for the viability of bacteria.

Similar to λ N, but with the opposite effect – termination instead of antitermination –, Nun protein of phage Hong Kong 022 binds to *boxB*, a stem-loop structure of the *nut* sequence. By *in vivo* studies the functional Nun Tyr39Ala mutation was identified to have important implications for this *boxB* interaction. The interaction of Tyr39 with A9 of the *boxB*, which was judged to be crucial earlier, is no longer possible with this mutation. Molecular dynamic calculations showed significant differences to the λ N:*boxB* interaction, where the analogous Trp18:A9 interaction is essential for processive antitermination. Thus this analysis gave important clues about the molecular switch between antitermination and termination.

1 Einleitung

1.1 Der Phage λ

Einblicke in komplexe zelluläre Vorgänge, wie z.B. Transkription, Translation und Signaltransduktion, auf molekularer Ebene sind eine zentrale Errungenschaft der Biochemie und Biophysik. Eines der am besten untersuchten Modellsysteme für das Verständnis der Transkription ist der 1951 entdeckte Phage λ (Lederberg, 1951). Die Bedeutung der detaillierten Einblicke in die Transkription wurde durch die Verleihung des Nobelpreises für Chemie im Jahre 2006 an Roger D. Kornberg für seine Forschungen zur molekularen Basis der eukaryontischen Transkription eindrucksvoll verdeutlicht. Vor wenigen Jahren konnten verschiedene Kristallstrukturen des zentralen Proteins der Transkription, der RNA Polymerase (RNAP), gelöst werden: eine bakterielle RNAP aus *Thermus thermophilus* (Vassylyev *et al.*, 2002) und die eukaryontische RNAPII aus Hefe (Cramer *et al.*, 2001; Gnatt *et al.*, 2001). Neben diesen strukturellen Untersuchungen gibt es eine Vielzahl neuerer genetischer und biochemischer Erkenntnisse, die es ermöglichen, das komplizierte Netzwerk der Regulation der Transkription sukzessive zu entflechten und zu verstehen. Hierfür bietet der Phage λ ein ideales Modellsystem, da er das Transkriptionssystem seines Wirts *Escherichia coli* (*E. coli*) für seine eigene Expression mitnutzt. Dabei benutzt er zwei unterschiedliche Strategien; entweder λ lässt *E. coli* eine große Anzahl von Viruspartikeln produzieren und dann durch Zelllyse in die Umgebung abgeben (lytischer Zyklus) oder λ sorgt für die eigene Integration ins Wirtsgenom (lysogener Zyklus). In diesem Zustand besitzt er aber weiterhin die Möglichkeit wieder in den lytischen Zyklus zu wechseln (Gottesman und Weisberg, 2004; Court *et al.*, 2007).

1.2 Lytischer und lysogener Zyklus

Nach der Infektion einer *E. coli* Zelle mit dem Phagen λ kann das System entweder in den lytischen Zyklus oder den lysogenen Zyklus eintreten. Beim lytischen Zyklus werden mehrere hundert Viruspartikel in der Zelle produziert und schließlich durch Zelllyse an die Umgebung abgegeben (Little, 2005). Die Transkription startet an den rechten und linken Promotoren (p_R und p_L) des Phagenoms und wird bis zu den Terminatoren t_R bzw. t_L fortgesetzt (Abb. 1-1; (Court *et al.*, 2007)). Das zuerst

transkribierte λ N-Protein, kontrolliert durch das zelluläre Protein RNaseIII, sorgt für eine Modulation der RNA Polymerase (RNAP), die sie resistent für t_R und t_L macht (Friedman und Court, 2001). Dieser als Antitermination (vgl. Abschnitt 1.4.1) bezeichnete Prozess ermöglicht die Transkription der stromabwärts von t_R bzw. t_L gelegenen Gene. Dieser Bereich enthält unter anderem die für die Replikation verantwortlichen Gene (o , p , ren), die sowohl im lytischen als auch im lysogenen Zyklus essentiell sind. Zusätzlich wird auch das regulatorische Gen q exprimiert. Dessen Genprodukt λ Q bewirkt, ähnlich wie λ N, Antitermination (Roberts *et al.*, 1998). Sie findet am t_R -Terminator statt, was zur Expression der sogenannten späten Gene führt, deren Proteine für die Zelllyse verantwortlich sind.

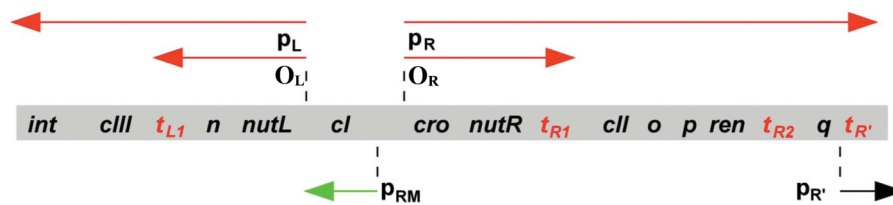


Abbildung 1-1: Ausschnitt aus der Genkarte des Phagen λ

Die frühen Transkripte der p_L und p_R Promotoren und deren Operatoren O_L bzw. O_R (Bindestellen für CI und Cro) sind als rote Pfeile gekennzeichnet. Das späte Transkript ab $p_{R'}$ ist als schwarzer Pfeil dargestellt. Gene sind schwarz und die Terminatoren sind rot hervorgehoben. 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 cl Gens stromabwärts von p_{RM} (grüner Pfeil) wichtig ist. Die Funktionen der anderen Gene sind im Text aufgeführt (nach Court *et al.*, 2007)).

Beim lysogenen Zyklus wird das Phagen genom in das Wirtsgenom integriert und die meisten Phagenproteine werden abgeschaltet, sodass λ in Form eines Prophagen vorliegt. Der Beginn des lysogenen Zyklus läuft identisch wie beim lytischen ab (Dodd *et al.*, 2005). Nach erfolgter λ N Expression und dem Überlesen der t_R bzw. t_L Signale werden die Gene int , cII und $cIII$ exprimiert (Abb. 1-1). Das int -Genprodukt ist eine Integrase die für den Einbau der Phagensequenz ins Wirtsgenom sorgt. Die CII- und CIII-Proteine aktivieren das cl -Gen, das für den λ -Repressor kodiert (Kaiser, 1957). Der λ -Repressor inhibiert durch Bindung an den O_R - bzw. O_L -Operator die Promotoren p_R und p_L (Atsumi und Little, 2006). Der Phage λ liegt nun als Prophage im Wirtsgenom vor. Durch äußere Aktivierung, z.B DNA-Schädigung, die zum Abbau

des λ -Repressors führt, kann der Prophage wieder in den lytischen Zyklus überführt werden (Abb. 1-2; (Ptashne, 2006)).

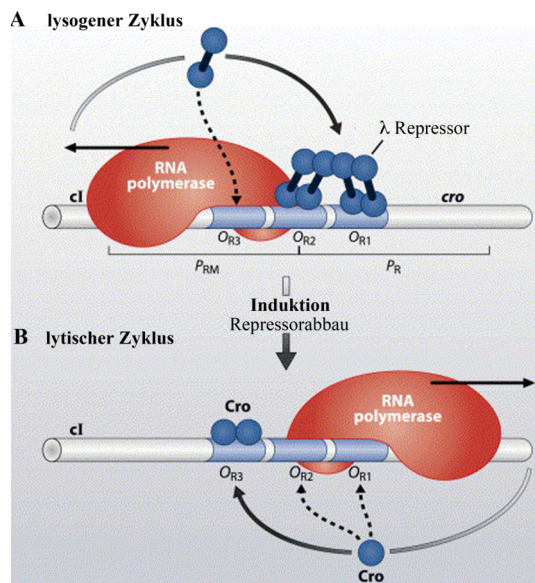


Abbildung 1–2: Übergang des λ Prophagen in den lytischen Zyklus

A) Ein in das Wirtsgenom integrierter λ Prophage wird durch den λ -Repressor, Genprodukt von *cI*, an der Transkription seiner lytischen Gene gehindert. Der Abbau des Repressors geschieht durch Aktivierung des RecA-Proteins (DNA-Reparatur-Protein) und sorgt für den Übergang des Prophagen in den lytischen Zyklus.

B) Durch die Transkription der lytischen Gene wird das *cro*-Genprodukt exprimiert, das seinerseits die freigewordenen Bindestellen des *cI*-Genprodukts besetzt und so eine Rückkehr zum lysogenen Zyklus unterbindet (Ptashne, 2006).

1.3 Bakterielle Transkription

1.3.1 Die bakterielle RNA-Polymerase

Die RNA-Polymerase (RNAP) ist das zentrale Enzym für Genexpression und Genregulation in allen drei Gattungen der Lebewesen: Eukaryoten, Archaeen und Bakterien (Murakami und Darst, 2003; Borukhov und Nudler, 2008). Das Enzym synthetisiert eine RNA-Kopie des codierenden DNA-Strangs aus Ribonukleosidtriphosphaten (NTP). Die RNAP verwendet chemische Energie (Wang *et al.*, 1998; Galburt *et al.*, 2007), um sich entlang der DNA gerichtet mit einer Geschwindigkeit von 15-80 (nt)/s (Nukleotideinbau pro Sekunde) bewegen zu können (Shilatifard *et al.*, 2003; Epshtein und Nudler, 2003). Diese Bewegung erfolgt jedoch diskontinuierlich und wird als Ratschen-Bewegung beschrieben (Bar-Nahum *et al.*, 2005; Abbondanzieri *et al.*, 2005; Brückner und Cramer, 2008). Diese diskontinuierliche Bewegung konnte für das Ribosom bei der Translation bereits früher gezeigt werden (Frank und Agrawal, 2000; Frank *et al.*, 2007). Strukturell sind alle RNA-Polymerasen aus verschiedenen Untereinheiten mit einer Gesamtgröße von mehreren hundert Kilodalton (kDa) aufgebaut, wobei im Gegensatz zu den Eukaryoten die Prokaryoten nur eine einzelne RNAP besitzen. Obwohl die Anzahl der Unter-

einheiten zwischen den Spezies variiert, ist ihre Sequenz, Tertiärstruktur, wie auch der Aufbau des katalytischen Zentrums evolutionär konserviert (Ebright, 2000; Minakhin *et al.*, 2001).

Die *E. coli* RNAP besitzt eine Molekülmasse von etwa 400 kDa, und hat einen α_2 , β' , β , ω -Untereinheitenaufbau, der auch als *core*-RNAP bezeichnet wird (Severinov, 2000). Die β - und β' -Untereinheiten bilden das aktive Zentrum und katalysieren somit die RNA-Polymerisation. Die amino-terminalen Domänen der α -Untereinheiten (α NTD) sind für den korrekten Zusammenbau der RNAP verantwortlich; die über einen flexiblen Linker verbundenen carboxy-terminalen Domänen (α CTD; (Jeon *et al.*, 1997)) haben regulatorischen Einfluss auf die RNAP durch unterschiedliche Wechselwirkungen mit Regulator-Proteinen (Benoff *et al.*, 2002). Für die ω -Untereinheit konnte gezeigt werden, dass sie nicht essentiell für die Funktionalität der RNAP ist, jedoch unterstützend auf den Zusammenbau der RNAP wirkt (Mukherjee und Chatterji, 1997; Mukherjee *et al.*, 1999; Mathew und Chatterji, 2006).

In Abbildung 1-3 ist die Kristallstruktur der RNAP aus *Thermus thermophilus* in der Elongationsphase dargestellt. Die einzelnen Untereinheiten sind gut zu erkennen, wobei die α CTD nicht gezeigt ist, da aufgrund ihrer Flexibilität keine Elektronendichte für sie bestimmt werden konnte (Vassylyev *et al.*, 2007a). Die RNAP besitzt drei Bindestellen, die Doppelstrang-Bindungsstelle (DBS), die Hybrid-Bindungsstelle (HBS) und die RNA-Bindungsstelle (RBS). Vom Aufbau gleicht sie einer Krebsschere, die mit ihren Zangen das DNA-Duplex (mit der DNA-Klemme) und die naszierende RNA (mit der aus β -Klappe, β' -Zink-Finger, β' -Verschluss und β' -Deckel zusammengesetzten RBS) stabilisiert. Gleichzeitig bildet sie in ihrem Inneren, flankiert von der β - und der β' -Untereinheit, den Raum für die Transkriptionsblase aus (Abb. 1-3).

Die DBS bindet ~ 9 bp des DNA-Doppelstrangs flussabwärts (*downstream*) des aktiven Zentrums (Nudler *et al.*, 1996; Nudler *et al.*, 1998; Korzheva *et al.*, 2000). Die DNA behält in diesem Bereich ihre typische B-Form bei und wird zum aktiven Zentrum hin um 90° geknickt, was das Aufschmelzen des Doppelstrangs ermöglicht (Gnatt *et al.*, 2001; Kettenberger *et al.*, 2004; Vassylyev *et al.*, 2007a).

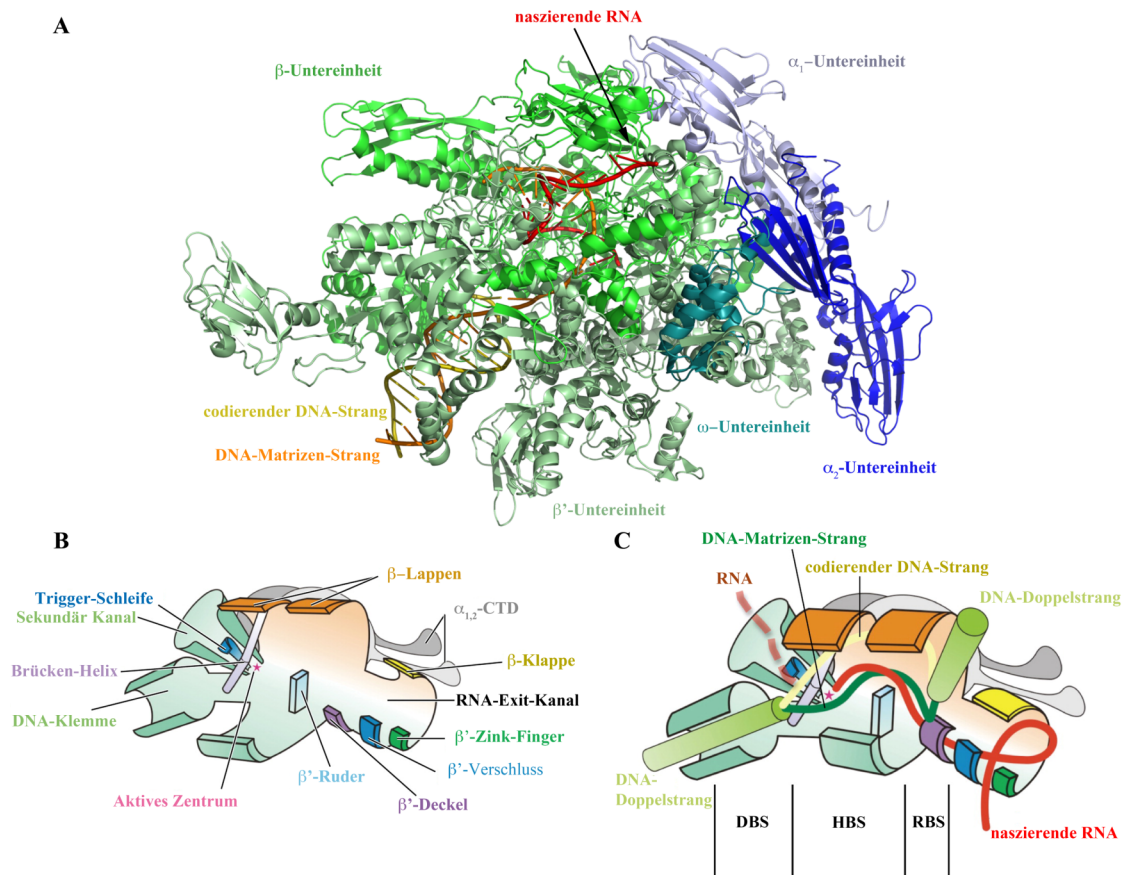


Abbildung 1-3: Struktur der bakteriellen RNA Polymerase

A) Kristallstruktur der RNAP aus *T. thermophilus* (PDB: 2O5I; (Vassylyev *et al.*, 2007a).
 B) Schematische Repräsentation der α -, β - und β' -Untereinheiten der RNAP Struktur als vereinfachtes Modell mit allen wichtigen strukturellen Merkmalen (Epshtein *et al.*, 2007). Die Funktionen der Strukturmerkmale sind im Text aufgeführt.
 C) Schematische Darstellung wie in B) einer aktiven RNAP. Im Primär Kanal wird das RNA-DNA-Hybrid in der Hybrid-Bindungsstelle (HBS) geschützt. Die β' -Domäne klemmt die flussabwärts liegende DNA in der DNA-Duplex-Bindungsstelle (DBS) ein. Der RNA-Exit-Kanal wird von β' -Verschluss, β' -Zink-Finger, β' -Deckel und β -Klappe gebildet. Die RNA-Bindungsstelle (RBS) ist an der Abzweigung der HBS und des RNA-Exit-Kanals und sorgt durch das β' -Ruder wie ein Reißverschluss für die Trennung der naszierenden RNA vom RNA-DNA-Hybrid (Nudler, 2009).

Das katalytische Zentrum wird von zwei Mg^{2+} -Ionen gebildet, wobei das erste Ion von drei Aspartatresten des konservierten NADFDGD-Motiv der β' -Untereinheit koordiniert wird (Mustaev *et al.*, 1997; Zhang *et al.*, 1999). Das zweite Mg^{2+} ist weniger stabil über einen Aspartatrest und vermutlich zwei H_2O -Moleküle koordiniert (Vassylyev *et al.*, 2007b). Die Mg^{2+} -Ionen katalysieren zusammen mit Teilen der β - und der β' -Untereinheit (Brücken-Helix und Trigger-Schleife) die Bildung der Phosphodiesterbindung zwischen der freien 3'OH-Gruppe der bereits vorhandenen RNA und dem nächsten einzubauenden NTP. Das erste Mg^{2+} aktiviert die freie OH-

Gruppe und das zweite Mg^{2+} stabilisiert die Phosphatgruppen des NTPs (Wang *et al.*, 2006; Vassylyev *et al.*, 2007b). Die Größe der Transkriptionsblase wird durch die β -Lappen reguliert. Sie verschließen die HBS und sorgen so dafür, dass das RNA-DNA-Hybrid eine Länge von 7-10 bp nicht übersteigt (Nudler *et al.*, 1997; Korzheva *et al.*, 2000). Zusätzlich wird das Hybrid durch das β -Ruder stabilisiert, um die Re-Assoziation der RNA mit der DNA flussaufwärts der Transkriptionsblase zu verhindern (Kuznedelov *et al.*, 2002; Westover *et al.*, 2004; Vassylyev *et al.*, 2007a).

Die RBS stabilisiert die naszierende RNA und unterstützt so zusätzlich die Gesamtstabilität des Komplexes. Der β' -Deckel formt zusammen mit Teilen der β -Untereinheit (auch β -Sattel genannt) eine enge Pore, durch welche die einzelsträngige RNA hindurchgefädelt wird (Gnatt *et al.*, 2001; Kettenberger *et al.*, 2004; Vassylyev *et al.*, 2007a). Im weiteren Verlauf weitet sich diese Pore zum RNA-Exit-Kanal, der von β' -Verschluss, β' -Zink-Finger und β -Klappe gebildet wird (Nudler *et al.*, 1997; Korzheva *et al.*, 2000), wobei letztere den Austrittskanal bedeckt und ein mögliches Ziel verschiedener regulatorischer Signale darstellt, die in die Termination involviert sind (Toulokhonov und Landick, 2003).

Der Nukleotideinbauzyklus lässt sich in drei Phasen unterteilen: Substratzugang und Bindung an den Elongationskomplex (NTP-Bindung), Nukleotideinbau und Translokation in Bezug auf die RNAP zum erneuten Freisetzen der NTP-Einbau-Position (A oder i+1 Position; Abb. 1-4; (Borukhov und Nudler, 2008; Nudler, 2009)). NTPs gelangen höchstwahrscheinlich über den sekundären Austrittskanal in die RNAP (Batada *et al.*, 2004). Sie werden gebunden und von den beiden Mg^{2+} -Ionen an der Trigger-Schleife aktiviert. Diese Bindung wird zusätzlich durch Leu1081 und His1085 der Trigger-Schleife unterstützt. Dadurch verschiebt sich das Gleichgewicht des freien Zustands zwischen prä- und posttransloziert (Nudler, 1999; Vassylyev *et al.*, 2007a; Brückner und Cramer, 2008) und induziert eine strukturelle Umfaltung der Trigger-Schleife zur Trigger-Helix. Dies sorgt für die korrekte Positionierung des NTPs für den Einbau in die RNA und den Verschluss des Sekundär-Kanals durch die Trigger-Helix (Abb. 1-4; (Wang *et al.*, 2006; Vassylyev *et al.*, 2007b)).

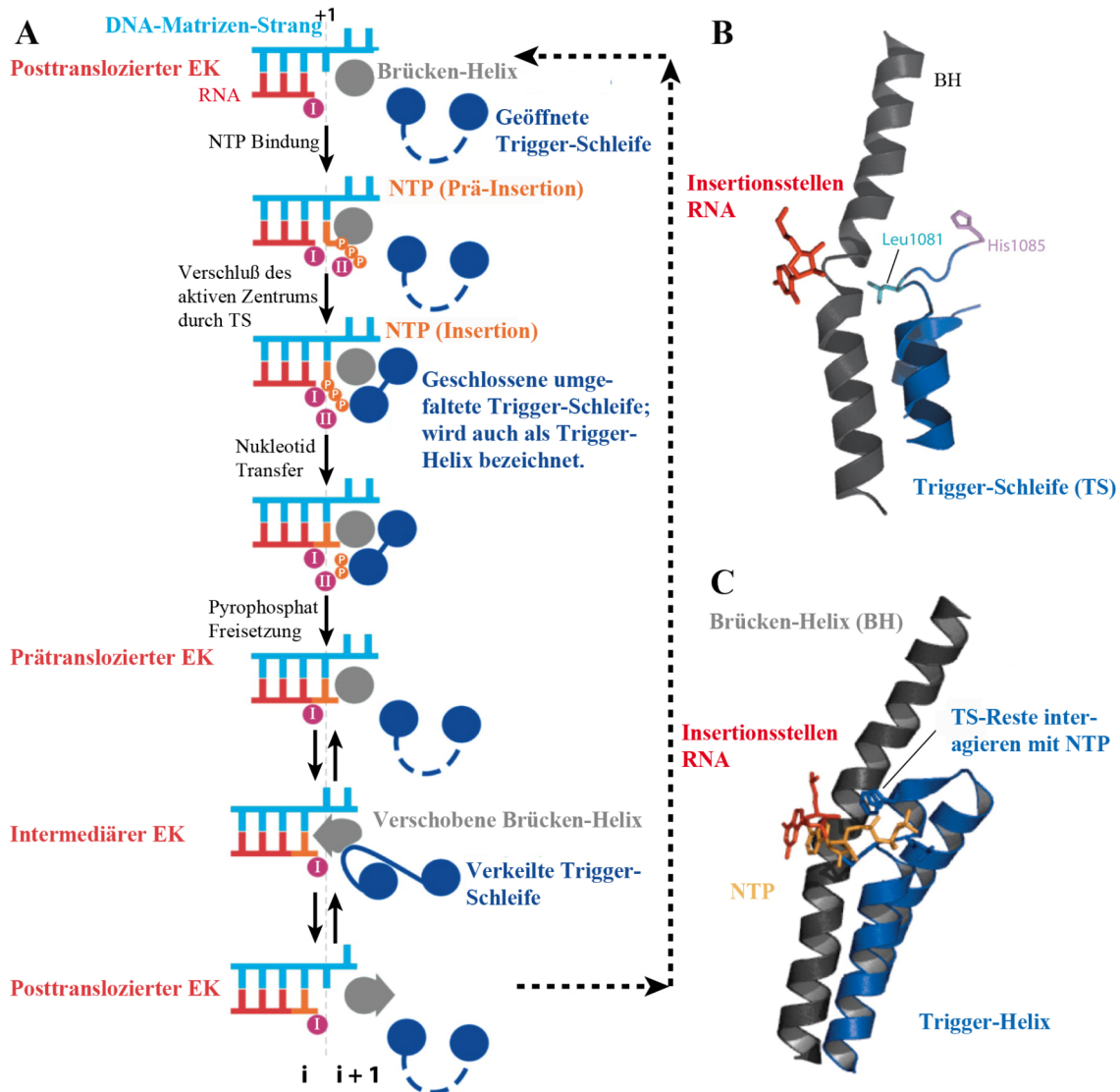


Abbildung 1–4: Nucleotideinbauzyklus und Dynamik des katalytischen Zentrums der RNA Polymerase (Nudler, 2009)

A) Schematische Darstellung der konformationellen Änderungen der Trigger-Schleife (TS) und der Brücken-Helix (BH) während des NTP-Einbaus und der darauf folgenden Translokation. Der codierende DNA-Strang ist in hellblau, der RNA-Strang in rot, die TS in dunkelblau und die BH in grau dargestellt. Magenta-Kreise markieren die katalytischen Mg^{2+} -Ionen. Am Beginn des Zyklus, posttranslozierter Elongationskomplex (EK), befindet sich TS in der ungefalteten geöffneten Konformation (B; (Wang *et al.*, 2006; Vassylyev *et al.*, 2007b)). Die Bindung des korrekten NTPs (orange) führt zur Umfaltung der TS in eine α -helikale Form (auch Trigger-Helix; C), diese Form stabilisiert ihrerseits mittels Leu1082 und His1085 die NTP-Bindung. Da die NTP-Bindung in der $i+1$ Position den EK in der posttranslozierten Form stabilisiert, dient das Substrat als Fixierung im Ratschen-Mechanismus der Translokation (Bar-Nahum *et al.*, 2005). Nach Einbau des NTPs in die RNA und Freisetzen des Pyrophosphats wechselt die TS/BH Einheit zwischen unterschiedlichen Konformationen, der geöffneten und der verkeilten, wobei letztere wohl zu einem Verschieben des RNA-DNA-Hybrids führt und somit die Translokation aktiv in der posttranslozierten Phase unterstützt. Basierend auf diesem Modell dient die TS/BH-Einheit als zweite wichtige Fixierung und treibt die Translokation in Abwesenheit eines Substrats voran (Gnatt *et al.*, 2001; Bar-Nahum *et al.*, 2005; Brückner und Cramer, 2008) und ein neuer Einbauzyklus kann gestartet werden.

Durch Ausbildung der Phosphodiesterbindung an der RNA wird die Trigger-Helix destabilisiert und verkeilt sich. Dies führt zu einem Verschieben der Brücken-Helix und somit letzten Endes zur Translokation auf die $i+1$ -Position (Bar-Nahum *et al.*, 2005; Brückner und Cramer, 2008). Nach erfolgter Translokation bildet sich die Trigger-Helix wieder zur Trigger-Schleife um und befindet sich wieder in der Ausgangskonformation (Komissarova und Kashlev, 1997a; Nudler, 1999; Vassylyev *et al.*, 2007a; Brückner und Cramer, 2008).

1.3.2 Initiation

Promotor spezifische Initiation der Transkription benötigt die Bindung des σ -Faktors an die *core*-RNAP, wodurch das Holo-Enzym gebildet wird (Helmann und Chamberlin, 1988; Burgess und Anthony, 2001). Dieses sucht entlang der DNA die entsprechenden Promotor-Erkennungssequenzen (Sakata-Sogawa und Shimamoto, 2004), bei denen es sich um hexamere Abschnitte konservierter Nukleotidsequenzen an den Positionen $-35'$ und $-10'$ der DNA handelt (Lisser und Margalit, 1993; Paget und Helmann, 2003). Nach erfolgter Promotorerkennung bildet sich die Transkriptionsblase durch Aufschmelzen des DNA-Doppelstrangs aus, wobei σ über verschiedene Reste und insbesondere ein konserviertes Threonin direkt daran beteiligt ist (Schroeder *et al.*, 2007; Schroeder *et al.*, 2008). Daraus resultiert ein stabiler Promotor-Komplex, der die RNA-Synthese startet. Die RNAP beginnt nun einen Zyklus, der als unvollständige Initiation bezeichnet wird; kurze RNA Abbruchfragmente werden synthetisiert, freigesetzt und gleich wieder neu synthetisiert (Hsu *et al.*, 2003). Neueste Untersuchungen deuten an, dass diese kurzen RNA-Stücke eine regulatorische Funktion haben könnten; sie könnten als Matrize für die Transkriptionsinitiation eines weiteren Promotors dienen oder als *Antisense*-Effektor für spezifische RNAs fungieren (Goldman *et al.*, 2009). Sobald die RNAP ein 8 bis 9 bp langes RNA-DNA-Hybrid gebildet hat, geht der Initiationskomplex in die Elongationsphase über (Abb. 1-5).

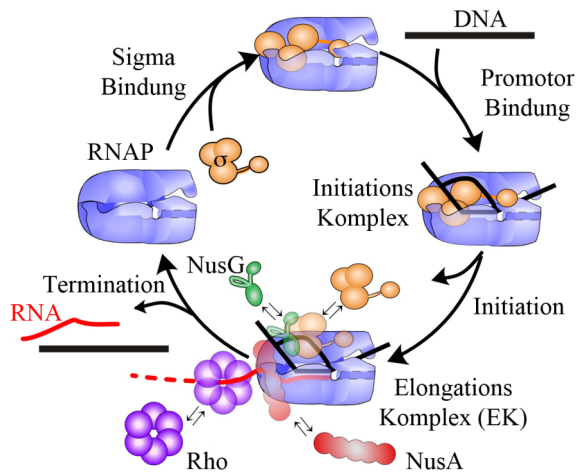


Abbildung 1–5: Schematische Darstellung der Initiation

Der σ -Faktor bindet an die RNAP und formt so das Holo-Enzym, das jetzt spezifisch an die Promotor-DNA binden kann und so die Transkription initiiert. Wenn das RNA-Transkript eine kritische Länge erreicht hat und das RNA-DNA-Hybrid dadurch stabilisiert ist, wird die Bindung zu σ gelockert, und ein stabilisierter EK wird ausgebildet. Der EK kann durch verschiedene Regulatoren, wie NusA, NusG und ρ , beeinflusst und modifiziert werden (Mooney *et al.*, 2009a).

In der Regel dissoziiert σ kurz nach dem Übergang in die Elongation von der RNAP (obligatorisches Modell (Metzger *et al.*, 1993)), dieses ist jedoch nicht zwingend notwendig für das Verlassen der Promotor-Region (Mooney *et al.*, 2005). Die Dissoziation von σ kann aber auch eher zufällig durch kontinuierliches Abnehmen der Affinität zur RNAP während der Elongation erfolgen (stochastisches Modell (Shimamoto *et al.*, 1986)), wobei sowohl *in vitro* als auch *in vivo* gezeigt werden konnte, dass ein kleiner Prozentsatz von σ -Faktoren an der RNAP assoziiert bleibt (Bar-Nahum *et al.*, 2005; Reppas *et al.*, 2006). Zusätzlich ist das NusG-Protein, das im weiteren Verlauf noch näher erläutert wird (Abschnitt 1.6.3), an der Dissoziation des σ -Faktors beteiligt, da es diesen von einer seiner Interaktionsflächen an der RNAP verdrängen kann (Mooney *et al.*, 2009a).

Die Familie der *E. coli* σ -Faktoren lässt sich in zwei Klassen einteilen: Klasse 1, die σ^{70} -Familie, umfasst die meisten σ -Faktoren, wie den σ^{70} -Faktor, der die Transkription der organisatorischen (*housekeeping*) Gene initiiert und somit auch kontrolliert (Borukhov und Severinov, 2002). Klasse 2, die σ^{54} -Familie, benötigt wie auch die eukaryontischen Initiationsfaktoren die ATP-Hydrolyse zur Energiegewinnung für das Aufschmelzen des DNA-Doppelstrangs (Buck *et al.*, 2000; Lin *et al.*, 2005). Funktionell ist diese Gruppe mit der σ^{70} -Familie verwandt, sie weist jedoch deutliche strukturelle Unterschiede auf (Paget und Helmann, 2003).

Die Erkennung der Promotoren durch σ -Faktoren erfolgt durch ihre DNA-Bindestelle. Im freien σ -Faktor ist diese Stelle durch eine andere Domäne blockiert, so dass die entsprechenden Promotoren nur in Gegenwart der RNAP gebunden werden können,

da die Autoinhibition erst durch Bildung des Holo-Enyms aufgehoben werden kann (Borukhov und Severinov, 2002).

Neben der Autoinhibition werden die verschiedenen σ -Faktoren auch über proteolytischen Verdau, Kontrolle der Neusynthese auf transkriptionaler und translationaler Ebene sowie Bindung an einen sogenannten anti- σ -Faktor reguliert, um eine kontrollierte Transkription zu gewährleisten (Hughes und Mathee, 1998; Helmann, 1999). Die Vielzahl der Regulationsmechanismen des σ -Faktors verdeutlicht seine Wichtigkeit bei der Initiation.

Die RNAP besitzt über ihre α CTD auch eine eigene wichtige regulatorische Funktion innerhalb der Initiation. Diese Regulation erfolgt auf zwei unterschiedliche Arten: Zum einen bindet die α CTD an eine stromaufwärts des Promotors gelegene *cis*-aktive DNA-Sequenz, das *upstream promotor* (UP)-Element (Estrem *et al.*, 1998; Estrem *et al.*, 1999). Durch diese Interaktion erhöht sich die Aktivität von Promotoren signifikant im Vergleich zum basalen Level (Ross *et al.*, 1993). Zum anderen konnte die Interaktion mit stromaufwärts gelegenen Transaktivatoren, wie CAP (*catabolite activator protein*) und OxyR (*oxydative stress regulator protein*) gezeigt werden (Zou *et al.*, 1992; Tao *et al.*, 1993; Benoff *et al.*, 2002).

1.3.3 Elongation

Mit der Bildung des RNA-DNA-Hybrids kommt es zur Umwandlung des Initiationskomplexes zum Elongationskomplex (EK). Dies führt zu signifikanten konformationellen Änderungen (Murakami und Darst, 2003), wobei dies auf die verringerte Affinität des σ -Faktors für die Promotor-DNA und die *core*-RNAP zurückzuführen ist (von Hippel und Pasmann, 2002). Generell hängt die Elongationsgeschwindigkeit von den Wachstumsbedingungen ab (Vogel und Jensen, 1994) und liegt in einem Bereich von etwa 40-90 Nukleotiden pro Sekunde (Gotta *et al.*, 1991; Condon *et al.*, 1993). Zusätzlich haben aber auch Antiterminationssequenzen, wie *boxA* in den ribosomalen RNA (*rrn*) Operons, einen direkten Einfluss auf die Geschwindigkeit (Vogel und Jensen, 1994). Neben dem korrekten Einbau der NTPs ist die RNAP auch in der Lage, diese Reaktion zur Korrektur falsch eingebauter Basen situativ durch Nuklease- oder Phosphorylase-Reaktionen umzukehren (Fish und Kane, 2002; Sosunov *et al.*, 2005; Brückner und Cramer, 2008; Nudler, 2009;).

Eine wichtige Regulation der Elongationsphase ist die Möglichkeit, die RNAP und somit auch die Transkription anzuhalten (Landick, 2006). Über diesen Mechanismus wird die Expression verschiedener Operons reguliert (Landick und Yanofsky, 1987; Donahue und Turnbough, 1994) und die Synchronisation von Transkription und Translation gesteuert (Winkler und Yanofsky, 1981; Yakhnin *et al.*, 2006). Der EK kann in einem angehaltenen, wartenden oder arretierten Zustand vorkommen.

In einem angehaltenen Zustand wird die RNAP aufgrund von Substratmangel versetzt. Dieser Zustand führt zu keinerlei strukturellen Änderungen innerhalb des Elongationskomplexes und sobald wieder genügend Substrat vorhanden ist, wird die Transkription fortgesetzt (Artsimovitch und Landick, 2000; Touloukhonov und Landick, 2003).

Ein wartender Zustand tritt aufgrund des Erkennens spezieller DNA-Signale, den Pausierungsstellen (*pause-sites*), ein (Gusarov und Nudler, 1999; Neuman *et al.*, 2003; Santangelo und Roberts, 2004). Dieser Zustand spielt eine wichtige Rolle sowohl bei der Kopplung von Transkription und Translation, als auch bei der Termination bzw. Antitermination. Für die Aufhebung dieses Zustandes sind keine zusätzlichen Faktoren notwendig. Es konnte für manche Pausierungsstellen sogar die Stabilisierung des wartenden Zustands durch Regulatorproteine gezeigt werden (Roberts *et al.*, 1998; Artsimovitch und Landick, 2002). Man unterscheidet die *pause-sites* in zwei Klassen: Klasse I zeichnet sich durch die Ausbildung einer Haarnadelstruktur der naszierenden RNA aus (Chan und Landick, 1993). Bei dieser Klasse beruht der Pausierungseffekt auf der Interaktion der Haarnadelstruktur mit den regulatorischen Untereinheiten des RNA-Exit-Kanals (vgl. Abb. 1-3), die zu einer Umordnung des aktiven Zentrums führt (Touloukhonov und Landick, 2003; Neuman *et al.*, 2003; Touloukhonov *et al.*, 2007). Klasse II ist durch ein schwaches RNA-DNA-Hybrid gekennzeichnet (Komissarova und Kashlev, 1997a; Komissarova und Kashlev, 1997b; Greive und von Hippel, 2005; Datta und von Hippel, 2008). Hier kommt es zu einer reversen Translokation und die 3'-liegenden Nukleotide werden vom RNA-DNA-Hybrid abgelöst. Die RNA-Kette wird in Richtung des sekundären Kanals verschoben. Es wurde angenommen, dass sie in diesen hineingeleitet wird, ihn so für neue Nukleotide versperren und auf diese Weise zum Pausieren der RNAP führt (Reeder und Hawley, 1996; Artsimovitch und Landick, 2000; Mejia *et al.*, 2008).

Neueste Untersuchungen zeigen jedoch, dass die revers translozierte RNAP genauso gut NTPs binden kann wie in ihrer aktiven Form. Dies deutet darauf hin, dass die reverse Translokation nicht der Hauptmechanismus dieser Art des Pausierens ist (Kireeva und Kashlev, 2009).

Aus diesem Pausieren mit reverser Translokation leitet sich der arretierte (*arrested* oder *dead-end*) Komplex ab, der nur durch die Zugabe von Proteinen, wie GreA oder GreB aufgehoben werden kann (Borukhov *et al.*, 1993; Kireeva *et al.*, 2005). Diese Protein-Faktoren binden mittels ihrer carboxy-terminalen Domänen in der Nähe des Sekundärkanals. Durch diesen gelangen sie *via* der amino-terminalen Domäne zum aktivem Zentrum der RNAP (Vassylyeva *et al.*, 2007). Dort hydrolysieren sie die Phosphatgruppen der 3'-liegenden Nukleotide, um die RNAP wieder in einen aktiven Zustand zu überführen (Koulich *et al.*, 1997; Laptenko *et al.*, 2003).

1.3.4 Termination

Neben den bereits angeführten Pausierungsstellen wird die Transkription an bestimmten Terminationssequenzen angehalten. Es kommt zu einer Destabilisierung des RNA-DNA-Hybrids. Dies führt wiederum zum Auflösen des EK, und die RNAP steht wieder für eine erneute Transkription zur Verfügung (Henkin, 1996; Henkin, 2000; Ciampi, 2006). Bei der Termination wird zwischen der intrinsischen und der Rho-abhängigen unterschieden, wobei beide Mechanismen erst nach einem vorherigen Pause-Zustand der RNAP aktiv werden können (Greive und von Hippel, 2005; Landick, 2006). Die Termination erfolgt nicht durch Blockierung der RNAP-Bewegung auf der DNA, sondern über die transkribierten RNA-Elemente (Martin und Tinoco, 1980; Lesnik *et al.*, 2001). Diese indirekte Termination führt dazu, dass Terminationssequenzen anhand der Effizienz des Signals und der Verweildauer der RNAP am Terminationssignal charakterisiert werden (Neuman *et al.*, 2003).

Für die intrinsische Termination werden keine zusätzlichen Proteine benötigt. Die Terminationssequenz besteht, ähnlich den *pause-sites*, aus einer Haarnadelschleife (Terminationsschleife), mit nachfolgender Uridin-reicher Sequenz. Dies ermöglicht die Schwächung des RNA-DNA-Hybrids (Lesnik *et al.*, 2001; Nudler und Gottesman, 2002; Nudler und Gusarov, 2003). Durch die Destabilisierung des Hybrids, wie auch durch die Haarnadelschleife, werden konformationelle Änderungen in der RNAP

induziert. Das Warten der RNAP erlaubt die korrekte Faltung der Terminationschleife (Gusarov und Nudler, 1999). Der detaillierte Mechanismus der intrinsischen Termination konnte bis *dato* noch nicht eindeutig geklärt werden. So existieren drei Modelle, die den EK während dieser Phase beschreiben sollen. Das im Augenblick präferierte allosterische Modell beruht auf der Interaktion der Terminationschleife mit strukturellen Elementen des RNA-Exit-Kanals, β -Klappe und β' -Zink-Finger. Dadurch werden Konformationsänderungen hervorgerufen, die das RNA-DNA-Hybrid weiter destabilisieren und zu einem Zusammenbrechen der Transkriptionsblase führen (Abb. 1-6; (Toulokhonov *et al.*, 2001; Epshtein *et al.*, 2007; Greive *et al.*, 2008; Larson *et al.*, 2008)).

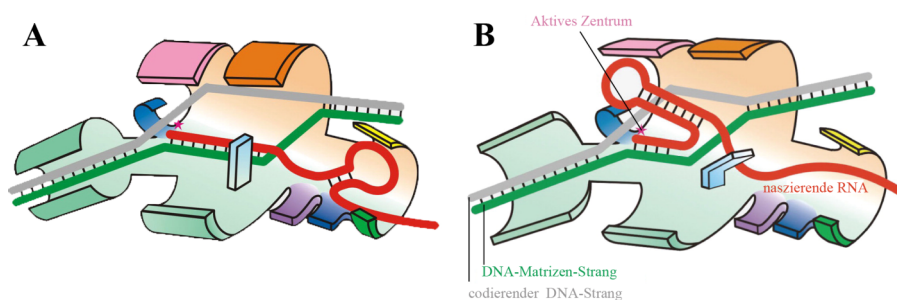


Abbildung 1–6: Allosterisches Modell der intrinsischen Termination

A) Die Ausbildung der Terminationschleife wird bei Pausierungsstellen durch β' -Zink-Finger und β -Klappe unterstützt und führt zu einer Aufweitung des RNA-Exit-Kanals.
 B) Die Haarnadelschleife hat ihre endgültige Größe von etwa 7-8 bp erreicht. Dadurch wird die Transkriptionsblase aufgeweitet, und die Terminationschleife interagiert mit der Trigger-Schleife. Dies führt zu einer Verkürzung des RNA-DNA-Hybrids auf etwa 3 bp. Zusätzlich wird durch die Haarnadelschleife die DNA-Klemme geöffnet, so dass neben der naszierenden RNA auch die DNA die RNAP verlassen kann (Epshtein und Nudler, 2003; Nudler, 2009).

Beim sogenannten RNA-Scher-Modell wird davon ausgegangen, dass das Herausziehen der RNA aus der Transkriptionsblase durch die Terminationschleife zum Auflösen des RNA-DNA-Hybrid führt und zusätzlich Konformationsänderungen der RNAP induziert (Macdonald *et al.*, 1993; Toulokhonov und Landick, 2003). Das dritte Modell nimmt an, dass durch die Terminationschleife die RNAP stromaufwärts geschoben wird, jedoch ohne eine entsprechende Transkription durchzuführen. Dies führt zum sogenannten 'hyper-translozierten' Stadium, das mit einem verkürzten RNA-DNA-Hybrid einhergeht und so den EK destabilisiert (Yarnell und Roberts, 1999; Santangelo und Roberts, 2004).

Die zweite Art der Termination ist vom ρ -Faktor (Roberts, 1969), einem homo-hexameren und ringförmig angeordneten Protein, abhängig. ρ ist essentiell für das

Überleben der Zellen, was auf seiner Rolle als globaler Regulator der Genexpression beruht (Cardinale *et al.*, 2008). Funktionell besitzt ρ eine ATP getriebene 5'-3'-RNA-DNA-Helikase, die es ihm ermöglicht, entlang der RNA in Richtung RNAP zu gelangen, das RNA-DNA-Hybrid aufzuschmelzen und so die Polymerase von der DNA abzulösen (Brennan *et al.*, 1987; Richardson, 2002). Diese Translokation wird durch NusG unterstützt, das ρ zur RNAP hin rekrutiert (Sullivan und Gottesman, 1992; Paman und von Hippel, 2000). Die Erkennungssequenz von ρ besteht aus der *rho-utilization (rut)* Sequenz, die entweder aus einer 30 nt langen cytidinreichen Sequenz oder einer Haarnadelschleife (Schneider *et al.*, 1993) sowie der stromabwärtsgelegenen Terminationssequenz besteht (70-80 nt), die jedoch lokal nicht eindeutig abgegrenzt ist (Alifano *et al.*, 1991; Banerjee *et al.*, 2006).

Die amino-terminale Domäne der ρ -Monomere enthält die primäre RNA-Bindungsstelle (Bogden *et al.*, 1999). Die carboxy-terminale Domäne besteht aus der P-Schleife, die für die ATP-Hydrolyse wichtig ist, der Q- und der R-Schleife, welche die sekundäre RNA-Bindestelle bilden (Burgess und Richardson, 2000; Burgess und Richardson, 2001). Durch die Strukturbestimmung mittels Röntgenkristallographie konnte gezeigt werden, dass die RNA schleifenartig entlang des Rings von den einzelnen Untereinheiten gebunden wird, und dass das 3'-Ende der RNA zur Ringmitte hin orientiert wird (Abb. 1-7A; (Skordalakes und Berger, 2003)). Interessanterweise öffnet sich der Hexamerring um etwa 12 Å nach dieser Primärbindung und gibt so die sekundären RNA Bindestellen im Inneren des Hexamers frei. Nun kann die RNA durch den Ring hindurchgefädelt werden, woraufhin der Ring wieder eine geschlossene Struktur einnimmt (Yu *et al.*, 2000; Kim und Patel, 2001; Skordalakes und Berger, 2003; Skordalakes und Berger, 2006). Angetrieben von der ATP-Hydrolyse kann sich ρ in dieser Form entlang der RNA zum EK hinbewegen, wobei der genaue Mechanismus hierfür noch nicht geklärt ist. Die eigentliche Termination, induziert durch das Aufschmelzen des RNA-DNA-Hybrids, erfolgt sehr wahrscheinlich wie es auch für die intrinsische Termination postuliert wurde (Abb. 1-7B; (Greive und von Hippel, 2005; Skordalakes und Berger, 2006)).

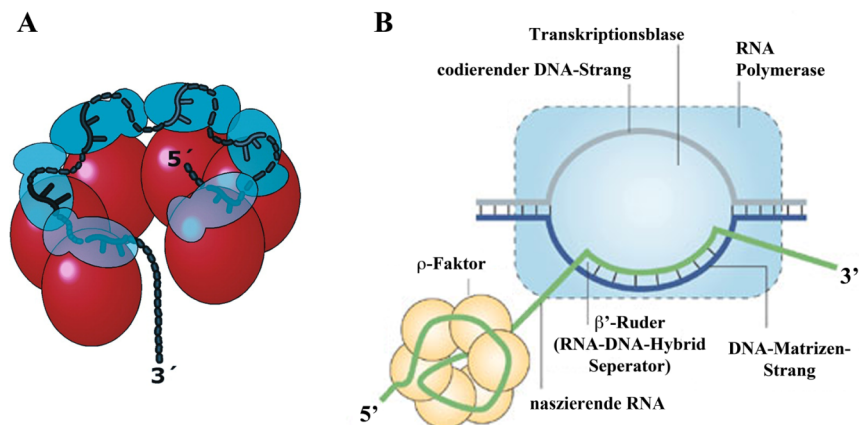


Abbildung 1–7: RNA-Bindung von ρ und Mechanismus der ρ -abhängige Termination

A) Schematische Bindung der RNA an die primären Bindungsstellen in der amino-terminalen Domäne von ρ (cyan). Die carboxy-terminalen Domänen sind in rot dargestellt und die RNA als schwarze Linie (Skordalakes und Berger, 2003).

B) Nach der ATP-getriebenen 5'-3'-Translokation von ρ entlang der RNA, kann ρ mit der RNAP interagieren. Termination wird durch die Destabilisierung des RNA-DNA-Hybrids induziert; wahrscheinlich durch strukturelle Änderungen innerhalb der RNAP und/oder aktive Verkürzung des Hybrids (nach (Greive und von Hippel, 2005)).

1.4 Antitermination

Als Antitermination wird das Überlesen von Terminationssequenzen, die generell nie vollständig terminieren können, und die daraus resultierende Expression von stromabwärts des Terminators gelegenen Genen bezeichnet. Erstmals beschrieben wurde dieser Effekt beim Phagen λ (Roberts, 1969). Generell besteht die Möglichkeit die Terminationsstellen positiv und negativ zu beeinflussen, wodurch eine weitere Regulationsmöglichkeit der Transkription gegeben ist (Weisberg und Gottesman, 1999). Die negative Regulation der Termination wird in zwei Arten unterschieden: Zum einen die Attenuation, die an *his* und *trp* Operons untersucht wurde und das Überlesen eines spezifischen Terminators ermöglicht (Henkin, 1996; Yanofsky *et al.*, 1996). Zum anderen die bereits erwähnte prozessive Antitermination. Durch die Interaktion mit einem Multi-Protein-Komplex ermöglicht sie es der RNAP nicht nur generell Terminationsstellen zu überlesen, sondern macht sie auch weniger sensitiv gegenüber Pausierungsstellen im Genom (Das, 1992; Friedman und Court, 1995; Roberts *et al.*, 2008). Neben dem Modellsystem des Phagen λ gibt es im Genom von *E. coli* auch ein eigenes Antiterminationssystem. Die Transkription der ribosomalen RNAs (*rrn*) unterliegt immer der regulatorischen Kontrolle durch Antitermination (Aksoy *et al.*, 1984; Condon *et al.*, 1995).

1.4.1 Das Antiterminationssystem des Phagen λ

Die Expression der frühen Gene bei λ wird durch das N-Protein reguliert, welches die Termination sowohl an intrinsischen als auch an ρ -abhängigen Terminatoren unterdrückt. Zusätzlich erhöht es auch die Transkriptionsrate (Mason *et al.*, 1992; Rees *et al.*, 1997). Beim 107 Aminosäure langen λ N handelt es sich um ein intrinsisch unstrukturiertes Protein mit folgenden Sequenzbereichen: Met1 – Asn22 binden die *boxB* der *nut*-Erkennungssequenz mit einer Affinität von $\sim 10^{-9}$ M, aufgrund der RNA-Bindung und der Sequenz gehört dieser Teil zur Gruppe der Arginin-reichen Motive (ARM; (Tan und Frankel, 1995; Cilley und Williamson, 1997; Van Gilst und von Hippel, 1997)). Der Bereich Asn34 – Arg47 interagiert mit der carboxy-terminalen Domäne von NusA (Abschnitt 1.5.1; (Bonin *et al.*, 2004; Prash *et al.*, 2006)) und der carboxy-terminale Teil Arg73 – Ser103 bildet Kontakte zur RNAP aus (Greenblatt *et al.*, 1998; Mogridge *et al.*, 1998).

Die Erkennungssequenz für die Antitermination von λ und somit zugleich auch der Interaktionspartner von λ N ist die als *N-utilization site (nut)* bezeichnete RNA-Sequenz, die zweimal im Phagengenom enthalten ist. Sie liegen jeweils stromaufwärts des p_L -Promotors (*nutL*) und des p_R -Promotors (*nutR*) und sind die ersten Terminatoren des Genoms (Friedman und Baron, 1974; Roberts *et al.*, 2008). Beide Sequenzen bestehen aus der einzelsträngigen *boxA* die über eine *Spacer*-Sequenz mit der *boxB* verbunden ist. *BoxB* ist eine Haarnadelschleife bestehend aus einem aus 5 Basenpaaren aufgebauten Stamm und einer aus 5 Basen gebildeten Schleife (Abb. 1-8A; (Hasan und Szybalski, 1986)). *BoxA* ist in beiden Sequenzen identisch und in der *boxB* ist in der *nutL* an der Position 9 der Schleife ein Guanin anstatt eines Adenins vorhanden. Die *Spacer*-Regionen zeigen deutliche Unterschiede (vgl. Abb. 1-8A).

Generell erfolgt die Regulation von λ N und somit auch der Antitermination bei λ mit drei unterschiedlichen Mechanismen. Zum einen kann die Synthese von λ N durch die Proteine Cro und CI gehemmt werden. Dies lässt sich auf die Blockierung des p_L Promotors zurückführen (Court *et al.*, 2007). Zum anderen kann die Aktivität des intrinsisch unstrukturierten N-Proteins über Proteaseabbau mittels verschiedener Proteasen, z.B. Lon, gesteuert werden (Maurizi, 1987; Court *et al.*, 2007). Und

drittens gibt es noch die Möglichkeit, dass sich λ N mit Hilfe der bakteriellen Endoribonuklease RNaseIII selber reguliert (Wilson *et al.*, 2002; Court *et al.*, 2007). In Abwesenheit von RNaseIII kann sich zwar der Antiterminationskomplex an der *nutL* ausbilden, jedoch ist die Kopplung mit der Translation durch eine große Stammschleifenstruktur gestört, die durch Blockieren der für das Ribosom notwendigen Shine-Dalgarno-Sequenz (SD) den Translationsstart inhibiert (Court *et al.*, 2007). Erst durch Abspalten der Haarnadelschleife, das durch die RNaseIII katalysiert wird, wird die SD freigesetzt, das Ribosom kann binden und die Translation startet (Wilson *et al.*, 2002; Wilson *et al.*, 2004).

Durch die Bindung an die 5'-Seite der *boxB* wird im arginin-reichen Motiv von λ N eine gebogene α -Helix induziert, die ihrerseits die Stammschleifen-Struktur der RNA stabilisiert. Diese Strukturmerkmale konnten mittels NMR-Spektroskopie gezeigt werden (Abb. 1-8B; (Legault *et al.*, 1998; Schärpf *et al.*, 2000). Die Stammschleife der *boxB* zeigt deutliche Strukturmerkmale der GNRA-Tetraschleifen, die sich durch eine uncharakteristische Basenpaarung zwischen dem ersten und vierten Nukleotid der Schleife und daraus resultierender hoher thermischer Stabilität auszeichnen (Jucker *et al.*, 1996). Bei der *boxB* wird durch Herausklappen der vierten Base der Schleife (Adenosin 9) erreicht, dass sich die übrigen vier Basen wie bei typischen GNRA-Tetraschleifen verhalten können. Es konnte gezeigt werden, dass die Bindung durch direkte Interaktionen im Stammbereich und insbesondere durch die π - π -Wechselwirkung zwischen Trp18 und A7 stabilisiert wird (Legault *et al.*, 1998; Schärpf *et al.*, 2000; Xia *et al.*, 2005).

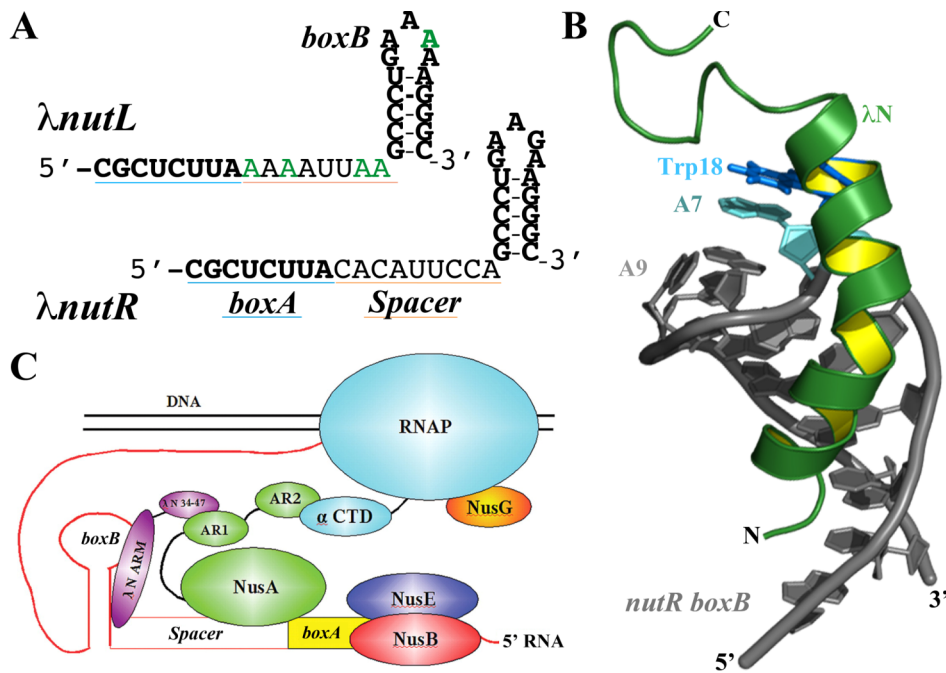


Abbildung 1–8: Der Antiterminationskomplex des Phagen λ

A) Die λ *nut* besteht aus der *boxA* (blau), der *Spacer*-Region (orange) und der Haarnadelschleife *boxB*, die eine GNRA-Tetraschleife bildet. Nukleotidaustausche zwischen λ *nutR* und λ *nutL* sind grün hervorgehoben (Lazinski *et al.*, 1989; Chattopadhyay *et al.*, 1995).

B) Das *arginine-rich motif* (ARM) des λ N Protein (grün) bindet als gekrümmte Helix an die 5'-Seite der *boxB* Haarnadelschleife. Diese Interaktion wird stabilisiert über die π - π -Wechselwirkung zwischen Trp18 (blau) und A7 (cyan; PDB: 1QFQ; (Legault *et al.*, 1998; Schärpf *et al.*, 2000)).

C) Der Komplex wird durch die Bindung der zusätzlichen Elongationsfaktoren NusA, NusB, NusE und NusG stabilisiert (Prasch *et al.*, 2009).

Diverse *in vivo* und *in vitro* Studien verdeutlichten, dass diese π - π -Wechselwirkung sowohl essentiell für die Stabilität des Komplexes wie auch die Effizienz der Antitermination ist (Franklin, 1993; Su *et al.*, 1997). Die Bedeutung dieser Trp18/A7-Wechselwirkung konnte zusätzlich durch *in vitro* Proteinselektion belegt werden (Xia *et al.*, 2003a). Dabei wurde festgestellt, dass die eine Helixwindung von Trp18 entfernten Aminosäuren 14 und 15, für die Bindungsaffinität entscheidend sind (Xia *et al.*, 2003a; Xia *et al.*, 2003b; Xia *et al.*, 2005). Interessanterweise konnte mittels Femtosekundenspektroskopie gezeigt werden, dass zwischen Trp18 und A7 ein Gleichgewicht zwischen offener (Verlust der π - π -Wechselwirkung) und geschlossener Konformation für die Reste Ala12 – Arg22 besteht (Xia *et al.*, 2003a; Xia *et al.*, 2005). Ohne die korrekte Ausbildung dieser π - π -Wechselwirkung verliert λ N seine Möglichkeit zur Antitermination (Xia *et al.*, 2005).

λ N kann alleine *in vitro* und *in vivo* die Transkription modifizieren (Rees *et al.*, 1996; Gusarov und Nudler, 2001). Der EK ist jedoch instabil, so dass nur Terminatoren in direkter Nähe zum Promotor überlesen werden können. Aus diesem Grund benötigt λ N, neben der *nut* und der RNAP, die *E.coli* Nus-Elongationsfaktoren (A, B, E, G; Abschnitt 1.6) für die Bildung eines stabilen Antiterminationskomplexes (Abb. 1-8C; (Mogridge *et al.*, 1998; Roberts *et al.*, 2008)). Diese Proteine beeinflussen und regulieren die RNAP so, dass sie die Terminationssequenzen nicht mehr erkennt und prozessive Elongation erfolgen kann (Mason und Greenblatt, 1991; Li *et al.*, 1993; Ciampi, 2006). Die *nut* Sequenz dient als Rekrutierungsstelle für die einzelnen Faktoren. Auf diese Weise wird die lokale Konzentration der Komponenten in der Nähe der RNAP erhöht (Mason und Greenblatt, 1991; Das *et al.*, 1996; Greive und von Hippel, 2005). Diese vielfältigen Protein-Protein und Protein-RNA Interaktionen stabilisieren den EK (vgl. Abschnitt 1.6). Dies resultiert in einem kürzeren Verweilen an Terminationstellen und somit einer Reduktion der Terminationseffizienz (Gusarov und Nudler, 2001; Nudler und Gottesman, 2002).

1.4.2 Die ribosomale (*rrn*) Antitermination

Neben der beschriebenen Antitermination des Phagen λ besitzt *E. coli* auch ein intrinsisches Antiterminationssystem. Die sieben verschiedenen ribosomalen RNA (*rrn*) Operons, *rrnA-E* und *rrnG-H* (Ellwood und Nomura, 1982), sind generell für die Anpassung der Wachstumsbedingungen notwendig (Condon *et al.*, 1995), ohne dass ihnen bis *dato* eine spezifische Funktion zugeordnet werden konnte (Condon *et al.*, 1992). In ihrem Grundaufbau sind alle sieben Operons gleich: Beginnend mit einer Leitsequenz folgt das 16S-rRNA-Gen, an das sich eine Trennsequenz anschließt, die operonspezifisch unterschiedliche tRNA-Gene enthält. Durch diese Trennsequenz werden die 16S-rRNA-Gene von den 23S-rRNA-Genen getrennt. Vor dem Terminator liegen die 5S-rRNA-Gene. Die Leitsequenz enthält unterschiedliche regulatorische Elemente: zwei σ^{70} -Promotoren, zwei Aktivatorsequenzen und ein Antiterminationssignal (Morgan, 1986; Condon *et al.*, 1995).

Aufgrund der hohen Sequenzhomologie zur *nut* konnten sowohl in der Leit- als auch in der Trennsequenz ein Antiterminationssignal bestehend aus *boxA*, *boxB* und *boxC* gefunden werden (Li *et al.*, 1984; Quan *et al.*, 2005). Allerdings liegen *boxA* und *boxB* in umgekehrter Reihenfolge auf dem Genom, wobei die *boxB*-Sequenz nicht

konserviert ist und für die Antitermination nicht essentiell ist (Gourse *et al.*, 1986; Berg *et al.*, 1989). Die starke Homologie der *boxA* deutet auf eine ähnliche Zusammensetzung des *rrn*-Antiterminationssystems hin wie beim Phagen λ (Abb. 1-9; (Squires *et al.*, 1993; Liu und Hanna, 1995; Greive *et al.*, 2005). Zusätzlich zu den Nus-Faktoren sind auch die ribosomalen Proteine L1, L3, L4, L13, S2 und S4 an der Transkriptionsregulation beteiligt (Torres *et al.*, 2001; Torres *et al.*, 2004).

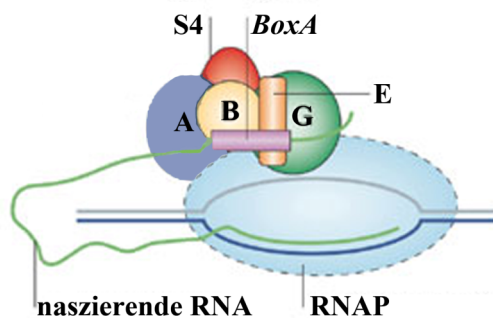


Abbildung 1-9: *rrn* Antiterminationskomplex

Der *rrn* Antiterminationskomplex besteht aus den gleichen Nus-Faktoren (A, B, E, G) und der *boxA* wie der λ Antiterminationskomplex. Zusätzlich interagiert das ribosomale Protein S4 mit dem Komplex und ersetzt das N-Protein und die *boxB* des λ Antiterminationskomplexes (modifiziert nach (Greive *et al.*, 2005)).

Im Gegensatz zur Antitermination bei λ kann das *rrn* Antiterminationssystem nur die ρ -abhängige Termination unterdrücken (Albrechtsen *et al.*, 1990; Condon *et al.*, 1995). Die *rrn*-Antitermination ist wichtig für die Regulation der Transkriptionsgeschwindigkeit, da eine langsamere Geschwindigkeit zum einen die korrekte Faltung der rRNA gewährleistet und es zum anderen ermöglicht Transkription und Translation zu koppeln (Winkler und Yanofsky, 1981; Squires *et al.*, 1993; Yakhnin *et al.*, 2006).

1.5 Die Termination des Phagen HK022

Der lambdoide Phage Hong-Kong 022 (HK022) ist vom funktionellen Aufbau des Genoms typisch für die λ Familie der temperenten Bakteriophagen. Im Detail lassen sich jedoch signifikante Besonderheiten erkennen, die im Fall von HK022 zu einem biologischen Vorteil gegenüber dem Phagen λ führen. Dieser Vorteil besteht in der Verhinderung einer Superinfektion, Infektion mit mehreren unterschiedlichen Bakteriophagen, von *E. coli*. Diese Superinfektion wird durch das 13 kDa große Nun-Protein verhindert, das ebenso ein arginin-reiches Motiv besitzt, um RNA zu binden. Das Nun-Protein konkurriert mit dem λ N-Protein um die Bindung an die *boxB* (Van Gilst *et al.*, 1997; Schärpf *et al.*, 2000; Faber *et al.*, 2001) und bindet in gleicher Art und Weise an die Stammschleifen-Struktur (Faber *et al.*, 2001). Zusätzlich besitzt Nun 20 weitere amino-terminale Reste, die regulatorische Aufgaben erfüllen können

(Stuart *et al.*, 2003). Die Bindung an die *boxB* ermöglicht es Nun, den EK von Anti-termination auf Termination umzuschalten (Oberto *et al.*, 1989). Dabei benutzt es neben der *nut* auch die Nus-Faktoren (Robledo *et al.*, 1991; Nudler und Gottesman, 2002). Die gegensätzlichen Funktionen von Nun und N können deshalb nicht nur auf die Interaktion mit der *boxB* zurückgeführt werden. Auch konnte gezeigt werden, dass der Carboxy-Terminus ein Zink-Bindemotiv und ein von der Position ungewöhnliches carboxy-terminales Tryptophan besitzt (Watnick und Gottesman, 1998; Watnick *et al.*, 2000). Die Koordination des Zinks erfolgt in Kombination mit der RNAP am β' -Zink-Finger durch Verdrängen von zwei Wassermolekülen. Zusätzlich interkaliert das Trp108 in die DNA-Doppelhelix und führt so zu einem arretierten Komplex (Watnick und Gottesman, 1999; Watnick *et al.*, 2000), der nicht mehr in der Lage ist die Elongation fortzuführen (Abb. 1-10).

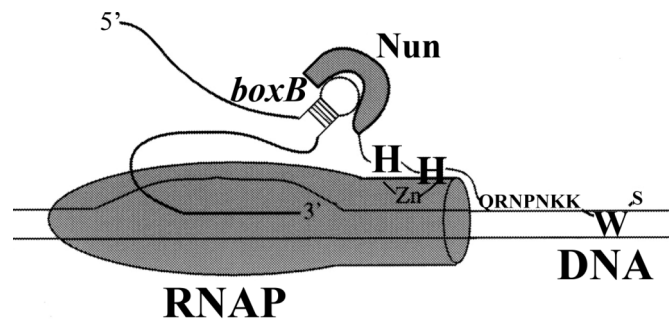


Abbildung 1–10: Modell der Nun-Termination

Nun bindet an die RNAP über ein vom β' -Zink-Finger koordiniertes Zn^{2+} -Ion. His98 und His93/His100 verdrängen dabei zwei Zn^{2+} koordinierende H_2O -Moleküle. Die basischen carboxy-terminalen Reste interagieren mit der RNA und Trp108 interkaliert in den DNA-Doppelstrang. Dies blockiert die Translokation und überführt die RNAP in einen arretierten Komplex (Watnick *et al.*, 2000)

1.6 Die Nus-Elongationsfaktoren

Die Nus-Faktoren wurden ursprünglich innerhalb des λ Antiterminationssystems identifiziert (Friedman und Baron, 1974). Daher leitet sich auch ihr Name *N-utilization substance* ab. Diese bakteriellen Faktoren stabilisieren den gesamten Elongationskomplex und führen so zu prozessiver Elongation (Mason und Greenblatt, 1991; Li *et al.*, 1993; Ciampi, 2006). Zusätzlich haben sie eine ähnliche stabilisierende Funktion bei der ribosomalen Antitermination an *rrn*-Operons (Li *et al.*, 1984; Quan *et al.*, 2005) und der vom Phagen HK022 induzierten Termination (Robert *et al.*, 1987).

1.6.1 NusA

Bei NusA handelt es sich um einen hochkonservierten bakteriellen Elongationsfaktor, der in Bakterien und Archaeen identifiziert werden konnte (Nudler und Gottesman, 2002). Dieses aus 495 Aminosäuren bestehende Protein ist in 6 Domänen organisiert: Die amino-terminale Domäne, für die eine direkte RNAP Interaktion postuliert wurde (Traviglia *et al.*, 1999; Borukhov *et al.*, 2005), ist über eine flexible Verbindungshelix mit dem zentralen Teil von NusA, bestehend aus der S1 und den zwei KH-Domänen, die kompakt und starr miteinander gekoppelt sind, verbunden (Gopal *et al.*, 2001; Worbs *et al.*, 2001). Für diesen Teil von NusA konnte die Interaktion mit naszierender RNA gezeigt werden (Arnvig *et al.*, 2004; Beuth *et al.*, 2005; Pagadala Santhanam, 2008; Prasch *et al.*, 2009). An diesen zentralen Teil schließen sich die bisher nur in *E. coli* gefundenen carboxy-terminalen Domänen AR1 und AR2 (*acidic repeat*) an (Abb. 1-11; (Eisenmann *et al.*, 2005)).

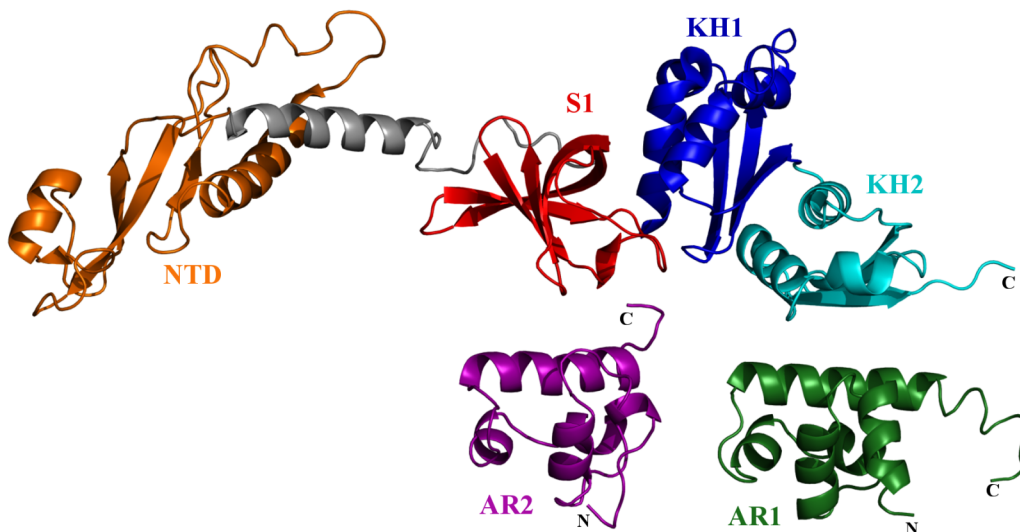


Abbildung 1–11: Struktur von *T. maritima* NusA und den beiden *E. coli* AR-Domänen

Die amino-terminale Domäne (NTD, orange) von NusA aus *Thermotoga maritima* ist über eine Verbindungshelix (grau) mit den drei RNA-Bindungs-Domänen S1 (rot), KH1 (blau) und KH2 (cyan) verbunden (PDB: 1HH2; (Worbs *et al.*, 2001)). Zusätzlich dargestellt sind die nur in *E. coli* vorkommenden Domänen AR1 (grün; PDB: 1WCL) und AR2 (violett; PDB: 1WCH; (Eisenmann *et al.*, 2005)).

Diese stark sauren Domänen sind in unterschiedliche Wechselwirkungen innerhalb der Transkription eingebunden. Für AR1 konnte die Bindung an das λ N Protein gezeigt werden (Bonin, 2004; Prasch *et al.*, 2006). AR2 hat zwei bis *dato* bekannte

zelluläre Funktionen. Es bildet mit der zentralen Domäne von NusA einen autoinhibitorischen Komplex indem es dessen RNA-Bindungsstellen maskiert (Mah *et al.*, 1999). Diese Autoinhibition wird durch die zweite Funktion von AR2 aufgehoben; die Bindung an die α CTD der RNAP (Mah *et al.*, 2000), die dazu führt, dass die α CTD nicht mehr an das UP-Element (*upstream promotor*) der DNA binden kann und somit der Elongationskomplex von der Initiation in die prozessive Elongation übergeht (vgl. Abschnitt 1.2.2; (Prasch, 2008)).

Generell kann für NusA gesagt werden, dass es primär die intrinsische Termination unterstützt und sobald Antiterminationssignale wie *nut* und *rrn* vorliegen, es die Antitermination fördert (Schmidt und Chamberlin, 1987; Gusarov und Nudler, 2001). Die intrinsische Termination beruht auf der Ausbildung der Terminationsschleife (vgl. 1.3.4; (Gusarov und Nudler, 1999)). Dieser Vorgang wird durch NusA beschleunigt. Dies lässt sich auf Interaktionen zwischen der S1-Domäne und der Haarnadelstruktur zurückführen (Gusarov und Nudler, 2001). Aufgrund von Western-Blot-Analysen wurde das Vorhandensein von zwei NusA Molekülen gebunden an die RNAP postuliert (Horwitz *et al.*, 1987). Auf diese Art und Weise können die gleichzeitigen Wechselwirkungen mit *nut* oder *rrn* und der Terminationsschleife erklärt werden (Gusarov und Nudler, 2001).

1.6.2 NusB und NusE (S10)

Homologe Proteine zum nahezu vollständig α -helikalen Protein NusB findet man in unterschiedlichsten Organismen (Altieri *et al.*, 2000). NusB bindet mittels seines amino-terminalen flexiblen arginin-reichen Motivs einzelsträngige RNA (Nodwell und Greenblatt, 1993; Lüttgen *et al.*, 2002; Mühlberger *et al.*, 2003): die *boxA* Sequenzen der *nut* und *rrn* Antiterminationssignale. Des Weiteren bildet NusB mit NusE einen stabilen Komplex, der diese RNA Affinität um den Faktor 10 steigert (Mason *et al.*, 1992; Nodwell und Greenblatt, 1993; Lüttgen *et al.*, 2002; Greive *et al.*, 2005). Diese Affinitätssteigerung beruht auf der Ausweitung der RNA-Bindungsfläche von NusB auf die Oberfläche von NusE (Luo *et al.*, 2008). Da NusE in freier Form keine stabile Struktur einzunehmen scheint (Culver und Noller, 1999; Lüttgen *et al.*, 2002; Greive *et al.*, 2005) und es außerdem ein elongiertes zweisträngiges β -Faltblatt besitzt, war die Strukturbestimmung erst im Komplex mit NusB und nach Entfernung dieses Faltblatts erfolgreich (Luo *et al.*, 2008). Strukturell führt die

Interaktion mit NusE nur zu einer leichten Umorientierung der α -Helices von NusB. NusE besteht aus vier zentralen β -Faltblättern, die auf einer Seite von zwei α -Helices flankiert werden (Luo *et al.*, 2008). Die Interaktionsfläche besteht auf NusE-Seite aus der Helix $\alpha 1$ und dem irregulären Faltblattstrang $\beta 2$ und auf NusB-Seite aus zwei helikalen Bündeln (Kontaktflächen I und II in Abb. 1-12). Diese mittels Röntgenkristallographie bestimmten Interaktionsflächen stimmen gut mit NMR-Daten über die Veränderung der chemischen Verschiebungen auf NusB-Seite überein (Das *et al.*, 2008).

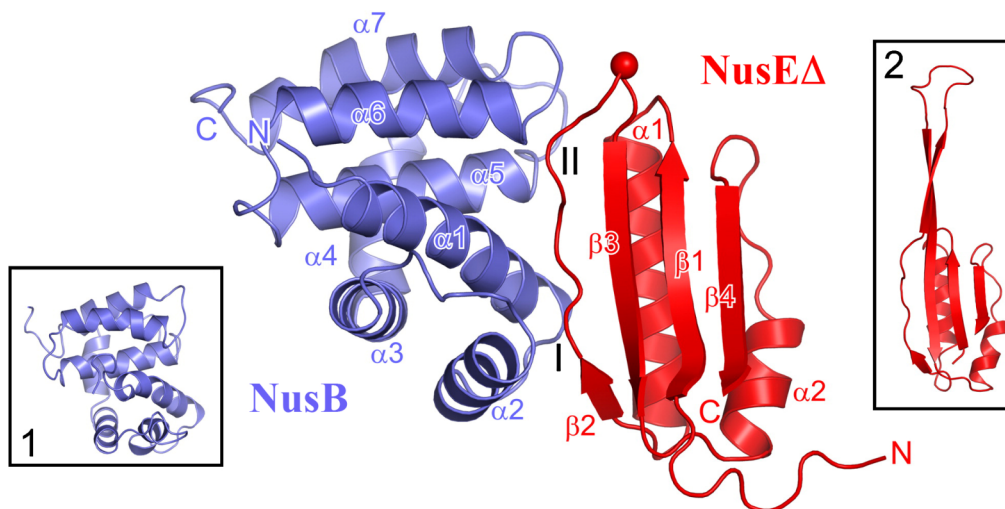


Abbildung 1–12: Struktur des NusB/NusE Δ Komplexes (Luo *et al.*, 2008)

Die Sekundärstrukturelemente und Termini von NusB (blau) und NusE Δ (rot) sind markiert (PDB: 3D3B). Die rote Kugel markiert die Position, an der das elongierte β -Faltblatt entfernt und durch ein einzelnes Serin ersetzt wurde. Interaktionsflächen auf NusB-Seite an der Seite des ersten drei-Helix-Bündels (I) und an der Spitze des zweiten drei-Helix-Bündels (II) sind hervorgehoben. 1: *E. coli* NusB NMR-Struktur (PDB: 1EY1 (Altieri *et al.*, 2000)) 2: Struktur von S10 aus der 30S ribosomalen Untereinheit (PDB: 2AVY (Schuwirth *et al.*, 2005)).

Durch die Studien von Luo *et al.* (2008) konnte gezeigt werden, dass NusE im NusB:NusE Komplex der aktive Partner ist (Weisberg 2008). NusB rekrutiert lediglich NusE zur *boxA*, was durch die Überexpression von NusE in einer NusB Deletionsmutante durch funktionale Antitermination belegt werden konnte (Luo *et al.*, 2008). Die Bedeutung von NusE wird durch seine essentielle Funktion verdeutlicht, da es neben der Transkription auch als S10-Protein der 30S ribosomalen Untereinheit direkt an der Translation beteiligt ist (Mizushima und Nomura, 1970; Friedman *et al.*, 1981; Wimberly *et al.*, 2000; Schluenzen *et al.*, 2000). Innerhalb der Translation gehört S10 zu den Proteinen, die für die Fertigstellung der 30S Untereinheit des Ribosoms verantwortlich sind (Noller und Nomura, 1996).

1.6.3 NusG

NusG ist ein essentieller Regulator der RNAP, der unterschiedlichste Effekte auf die Transkriptionsrate hat. NusG ist ein hochkonserviertes Protein, das man in allen Eubakterien findet und zu dem homologe Proteine in Eukaryonten (Hartzog *et al.*, 1998) und Archaeen (Kyrpides und Ouzounis, 1999) identifiziert werden konnten. *E. coli* NusG ist ein 181 Aminosäuren großer Elongationsfaktor bestehend aus zwei Domänen, die über einen flexiblen Bereich verbunden sind (Abb. 1-13; (Mooney *et al.*, 2009b)). Die amino-terminale Domäne (Aminosäuren 1-116) besteht aus vier zentralen anti-parallelen β -Strängen und drei umgebenden α -Helices (Abb. 1-13; PDB: 2K06). Die carboxy-terminale Domäne (123-181) ist aus fünf anti-parallelen β -Strängen aufgebaut, die sich in Form eines Fasses anordnen (Abb. 1-13; PDB: 2JVV). Die *E. coli* NusG Struktur ist fast identisch mit den korrespondierenden Domänen aus *Aquifex aeolicus* (Steiner *et al.*, 2002; Knowlton *et al.*, 2003) und *Thermus thermophilus* (Reay *et al.*, 2004), obwohl in beiden Organismen eine zusätzliche Domäne vorhanden ist, der bis jetzt noch keine Funktion zugeordnet werden konnte.

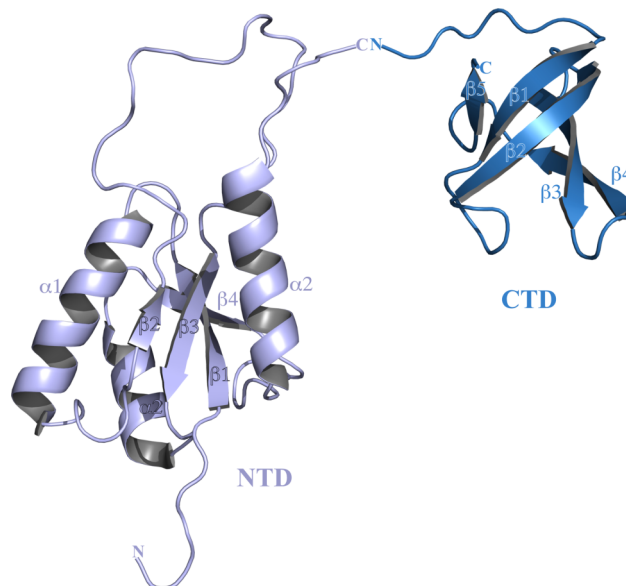


Abbildung 1–13: Struktur von *E. coli* NusG

NusG-NTD (hellblau; PDB: 2K06) und NusG-CTD (aquamarin; PDB: 2JVV) sind über einen flexiblen Linker miteinander verbunden und können so unterschiedliche regulatorische Funktionen ausüben (Mooney *et al.*, 2009b). Die Sekundärstrukturelemente und Termini sind markiert.

Funktionell ist NusG direkt an der ρ abhängigen Termination der Transkription beteiligt (Sullivan und Gottesman, 1992), indem es einerseits direkt mit dem ρ -Faktor interagiert (Li *et al.*, 1993; Paman und von Hippel, 2000) und andererseits die reverse Translokation der RNAP verhindert (Artsimovitch und Landick, 2000; Paman und von Hippel, 2000). Neben dieser Funktion ist NusG auch direkt in die Anti- bzw. Termination der lambdaoiden Phagen involviert (vgl. 1.3.4 und 1.4; (Sullivan und Gottesman, 1992; Mason *et al.*, 1992; Li *et al.*, 1992; Sullivan *et al.*, 1992; Burova *et al.*, 1999)). Die versatile Rolle von NusG bei der Transkriptionsregulation ist die Modulation des EK auf unterschiedliche Art und Weise. Es konnte sowohl *in vitro* als auch *in vivo* gezeigt werden, dass NusG die Elongationsrate der RNAP erhöht (Burova *et al.*, 1995; Burns *et al.*, 1998; Artsimovitch und Landick, 2000), was zumindestens teilweise auf das Unterdrücken von Transkriptionspausen zurückgeführt werden kann (Artsimovitch und Landick, 2000). Für diesen regulatorischen Effekt ist eine direkte Interaktion mit der RNAP verantwortlich (Mason und Greenblatt, 1991; Li *et al.*, 1992; Belogurov *et al.*, 2009). Die Regulation an den Klammern der β' -Untereinheit (*β' -clamp helices*) besteht für NusG darin, dass es an dieser Stelle den σ -Faktor von der RNAP verdrängt und somit den EK in die späte Elongationsphase überführt (Mooney *et al.*, 2009a). Für NusG konnte dieses erst kürzlich beschrieben werden (Belogurov *et al.*, 2009). Vom NusG Paralog RfaH (Belogurov *et al.*, 2007) ist der Mechanismus jedoch detaillierter bekannt. Es bindet mit seiner amino-terminalen Domäne, die sowohl sequentiell als auch strukturell hoch homolog zu NusG ist, an die β' -Klammern, über die Interaktion mit hydrophoben Resten, der RNAP (Sevostyanova *et al.*, 2008; Belogurov *et al.*, 2009). Die Spezialisierung des Paralogs RfaH besteht darin, dass es nicht wie NusG ein allgemeiner Elongationsfaktor ist, sondern erst durch die Bindung an das *ops*-Sequenzelement des DNA-Matrizen-Strangs aktiviert wird. Dadurch wird die Autoinhibition des Proteins, welche durch die Ausbildung einer geschlossenen Konformation durch direkte Interaktion der beiden Domänen erreicht wird, aufgehoben (Artsimovitch und Landick, 2002; Belogurov *et al.*, 2007). Diese geschlossene Konformation wurde auch für NusG aus *Aquifex aeolicus* unter bestimmten Kristallisationsbedingungen beschrieben (Knowlton *et al.*, 2003), jedoch konnte die biologische Bedeutung dieser Beobachtung noch nicht mit weiteren Untersuchungen bestätigt werden.

Über die genaue Art und Weise der beschriebenen Regulationen von NusG ist auf molekularer und struktureller Ebene wenig bekannt. Es wird diskutiert, dass die amino-terminale Domäne Nukleinsäuren bindet (Steiner *et al.*, 2002). In der carboxy-terminalen Domäne ist ein sogenanntes KOW-Motiv (Kyrpides *et al.*, 1996) vorhanden, das mit RNA-, Protein- und Ribosom-Interaktionen in Verbindung gebracht wird.

Zusätzlich konnte für NusG gezeigt werden, dass es direkten Einfluss auf die Geschwindigkeit der Translation hat (Zellars und Squires, 1999). Ein zusätzlicher Hinweis auf die Wichtigkeit dieser Regulation ist das Vorhandensein von etwa 10.000 – 20.000 NusG Molekülen (Li *et al.*, 1993) bei einer gleichzeitigen Anzahl von etwa 2.000 RNAP Molekülen in der Zelle (Squires und Zaporjets, 2000) unter optimalen Wachstumsbedingungen. Dieser signifikante NusG-Überschuss deutet auf eine substantielle Beteiligung von NusG an der Translation hin.

2 Ziele

Obwohl die Regulation der RNA-Polymerase und somit der Transkription immer besser verstanden wird, sind immer noch viele direkte Wechselwirkungen nicht identifiziert und der detaillierte Mechanismus ist weitgehend unbekannt. Die Antitermination des Phagen λ ist ein relativ gut verstandenes Modellsystem, für das eine solide Datenbasis vorhanden ist, um die regulatorischen Details der Transkription zu erfassen, zu analysieren und zu charakterisieren. Aus diesem Grund spielte insbesondere die Aufklärung der Rolle und Funktion von NusG in dieser Arbeit eine zentrale Bedeutung. Über NusG ist außer seiner kürzlich veröffentlichten Struktur (Mooney *et al.*, 2009b) nur gezeigt worden, dass es mit dem ρ -Faktor und der RNAP interagiert (Pasman und von Hippel, 2000; Belogurov *et al.*, 2009), ohne das jedoch strukturelle Details dieser Wechselwirkungen bekannt waren. Aufgrund der zwei Domänen-Struktur von NusG (Mooney *et al.*, 2009b) und seiner hohen Konzentration in der Zelle (Squires und Zaporozjets, 2000) wird angenommen, dass es eine wichtige Rekrutierungsfunktion im Antiterminationskomplex einnimmt. Die Domänen von NusG sollten zu diesem Zweck genauestens untersucht werden, um eventuelle Proteinbindungsflächen zu identifizieren. Im weiteren Verlauf sollten innerhalb des Antiterminationskomplexes mögliche Interaktionspartner analysiert werden, um Einblick in die Rolle und Funktion von NusG zu erhalten.

Des Weiteren war für NusB und NusE bekannt, dass sie ein Heterodimer bilden und mit der *boxA* wechselwirken (Mason *et al.*, 1992; Greive *et al.*, 2005). Diese Interaktion sollte detailliert mittels Fluoreszenzmessungen analysiert werden, um genaue Information über essentielle Bindesequenzen sowohl auf RNA- als auch auf Protein-Seite zu erhalten, wozu zusätzlich bekannte Mutationen auf beiden Seiten charakterisiert werden sollten.

Außerdem sollte der Effekt des Nun-Proteins auf den λ -Antiterminationskomplex untersucht werden, da bis auf Interaktionen des Carboxy-Terminus von Nun mit der RNAP und der DNA, die zu einem arretierten Elongationskomplex führen (Watnick *et al.*, 2000), die genauen Details des Umschaltens von Antitermination auf Termination noch nicht verstanden sind. Hier sollte insbesondere eine Mutation detailliert charakterisiert werden, die einen deutlichen Unterschied zum λ N-Protein

in der *boxB*-Bindung aufzeigt. Neben der experimentellen Analyse der Bindung wurden hierfür auch Molekular-Dynamik Berechnungen (MD) durchgeführt, um diese Effekte zu untersuchen. Die gewonnen Erfahrungen mit der MD sollten auch auf andere System wie die Faltung von Peptiden und die Stabilität von einzelnen Proteindomänen übertragen werden.

3 Zusammenfassung und Diskussion der Ergebnisse

3.1 Strukturelle Untersuchung von molekularen Komplexen

Zur strukturellen Untersuchung der zahlreichen Wechselwirkungen innerhalb des Antiterminationskomplexes auf atomarem Level bieten sich sowohl die Röntgenkristallographie als auch die NMR-Spektroskopie (*nuclear magnetic resonance*) an. In dieser Arbeit wurde die NMR-Spektroskopie angewandt, da sie, trotz der Größenlimitierung der untersuchten Proteine bzw. Komplexe, in Lösung abläuft und so wichtige Informationen über transiente Wechselwirkungen liefern kann. Generell ist die NMR-Strukturbestimmung von Proteinen und Proteinkomplexen bis zu einer Größe von 25 kDa mittels ^{13}C und ^{15}N Markierung heutzutage Routine. Mittlerweile ist es möglich durch zusätzliches Markieren der Proben mit ^2H und durch neuere NMR-Messtechniken, wie das TROSY (*transverse relaxation optimized spectroscopy*; (Pervushin *et al.*, 1997; Salzmann *et al.*, 1998)), das Problem der transversalen Relaxation, die für große Proteine zu einem mangelnden Magnetisierungstransfer führt (Clore und Gronenborn, 1998), zu umgehen. Durch einen modularen Ansatz, so dass beispielsweise einzelne Komponenten selektiv markiert werden können, sind auch deutlich größere Komplexe strukturell und dynamisch mittels NMR analysierbar (Fiaux *et al.*, 2002; Mittermaier und Kay, 2006).

Zur Untersuchung struktureller Änderungen bei Komplexbildung ist die Änderung der chemischen Verschiebung (*chemical shift perturbation* (CSP)), die im Fall einer Interaktion auf der Veränderung der Umgebung der beteiligten Kernspins beruht (Zuiderweg, 2002; Clarkson und Campbell, 2003), bestens geeignet. Diese Methode wurde für die Untersuchung der Interaktion zwischen NusG und NusE (Einzelarbeit F) angewandt, wobei in diesem Fall NusB in den Experimenten als Faltungshelfer für NusE benötigt wurde und somit aufgrund der Gesamtgröße von etwa 30-35 kDa die selektive Markierung mit ^2H für einige Experimente nötig machte. Außerdem wurde die mögliche Wechselwirkung zwischen den Domänen von NusG (Einzelarbeit E) mit dieser Methode untersucht. Es zeigte sich jedoch, dass diese Interaktion nur transient ist. Somit war der Einsatz einer Messmethode erforderlich, die von NMR Parametern bestimmt wird, die einen signifikanten Effekt zeigen, obwohl die entsprechenden Zustände niedrig populiert sind. Dies ist mit der Erhöhung der (transversalen)

Relaxation durch paramagnetischen Relaxations Verstärkung (*paramagnetic relaxation enhancement* (PRE)) möglich. Bei dieser Methode wird ein Radikal, z.B. Nitroxylradikal, an das Protein gekoppelt und der Einfluss der Kernspin-Elektronenspin Kopplung auf die NMR-Parameter untersucht. Diese Kopplung führt abstandsabhängig zu einer erheblichen Linienverbreiterung im NMR-Spektrum und lässt sich auch bei größeren Abständen ($< 30 \text{ \AA}$) zwischen Kernspin und Radikal beobachten (Clore *et al.*, 2007; Clore, 2008).

Da für NMR-Messungen Konzentrationen im höheren μM – mM -Bereich benötigt werden, können mittels NMR nur Dissoziationskonstanten in diesem Bereich, wie für die NusG:NusE Interaktion (Einzelarbeit F), quantifiziert werden. Für Dissoziationskonstanten im nanomolaren Bereich wurde deswegen auf die Fluoreszenz-Spektroskopie zurückgegriffen. Auf diese Weise wurden die Interaktionen zwischen NusB:NusE und RNA (Einzelarbeit D) sowie Nun-Peptiden und RNA (Einzelarbeit A) analysiert.

Für die Nun-RNA Interaktion wurden auch Molekular-Dynamik (MD) Berechnungen zur Stabilität der Peptid-RNA-Interaktion und zur Verifikation der experimentellen Daten durchgeführt. Diese Erfahrungen mit der MD wurden auch auf das System der humanen Guanylatcyclasen (hGC) zur Berechnung der Domänenstabilität von hGC-B (Einzelarbeit C) und der Faltung des Liganden STh der hGC-C (Einzelarbeit B) übertragen. Da sich bei dieser Thematik meine Beteiligung auf die MD-Berechnungen beschränkte und das Thema deutlich von meinem Hauptthema abweicht, werden die Einzelarbeiten B und C im weiteren Verlauf nicht detailliert diskutiert.

3.2 Die Bindung an die Schleife der *boxB* ist für Nun nicht essentiell

Es war bekannt und strukturell charakterisiert, dass sowohl λ N aus dem Phagen λ , als auch das Nun-Protein aus dem Phagen HK022 in gleicher Art und Weise an die *boxB* der *nut*-Sequenzen des λ Genoms binden (Legault *et al.*, 1998; Schärpf *et al.*, 2000; Faber *et al.*, 2001). Da durch das Nun-Protein im Gegensatz zu λ N jedoch Termination induziert wird, liefern Unterschiede in den jeweiligen Komplexen wichtige Hinweise auf den molekularen Schaltmechanismus. Es konnte eine Nun-Mutation, NunY39A, identifiziert werden, die erhebliche Effekte in *in vivo* Experimenten zeigt (Einzelarbeit A). Bei dieser Mutation ist eine als sehr wichtig

angesehene π - π -Interaktion zwischen Tyr39 und A9 der *boxB* nicht mehr möglich und führt so zu einer hohen Flexibilität in diesem Bereich (Abb. 3-1).

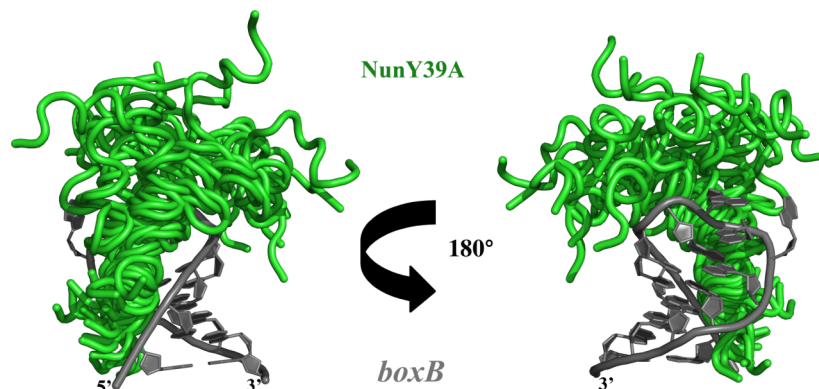


Abbildung 3-1: NunY39A-*boxB*-RNA Komplex

Überlagerung der mittels NMR bestimmten 20 Strukturen mit den niedrigsten Energien des NunY39A-*boxB*-RNA Komplexes. Die Peptidstruktur (grün) ist im Stammbereich der RNA (grau) gut definiert (α -helikal). Ab der Knickregion zwischen den Resten 31 und 33 zeigt sie ein hohes Maß an Flexibilität, ohne dass sie jedoch ihre Helizität verliert. Die Strukturbestimmung erfolgte mit einer Gesamtzahl von 189 intramolekularen und 13 intermolekularen Distanzbeschränkungen. In allen Strukturen sind die Distanzverletzungen $< 0.1 \text{ \AA}$ (Einzelarbeit A, Abbildung 6)

Die mittels NMR-Spektroskopie bestimmte Struktur (Abb. 3-1) konnte mit MD-Berechnungen belegt werden, wobei beide Methoden eine beträchtliche Flexibilität auf der ps – μ s Zeitskala zeigen. Bei Betrachtung des 1D-NMR-Spektrums konnten sowohl für Wildtyp Nun als auch für NunY39A die identische Stabilisierung der RNA anhand ihrer Iminosignale beobachtet werden (Einzelarbeit A). Dies deutete zum einen auf eine gleiche Bindung an die RNA hin und zum anderen darauf, dass die π - π -Wechselwirkung im Wildtyp keinen Einfluss auf die Struktur des Komplexes hat. Um dieses detaillierter zu untersuchen, wurden für beide Peptide die Dissoziationskonstanten mittels Fluoreszenzspektroskopie bestimmt. Hier zeigte sich für NunY39A nur eine marginale Erhöhung des K_d auf 6,1 nM gegenüber 2,4 für den Wildtyp. So formt sich aus den Daten ein einheitliches Bild: Die Wechselwirkung Nun-*boxB* ist von der π - π -Interaktion unabhängig, wobei *in vivo* Experimente die Terminationsfunktionalität dieser Mutante belegt haben (Einzelarbeit A). Dies ist eine wichtige Beobachtung, da für die gleiche π - π -Wechselwirkung zwischen Trp18 und A7 bei λ N gezeigt werden konnte, dass sie essentiell ist (vgl. Abschnitt 1.3.1; (Xia *et al.*, 2003a; Xia *et al.*, 2003b)). Dieser Punkt deutet auf zwei wichtige Aspekte hin: Zum einen ist es möglich, dass λ N für die Interaktion mit NusA fixiert werden muss

(Bonin *et al.*, 2004; Prash *et al.*, 2006) und dieses für die Nun-RNAP-Interaktion nicht notwendig ist (Watnick *et al.*, 2000). Zum anderen ist eine energetische Erklärung möglich, da der Antiterminationskomplex über einen deutlich längeren Zeitraum stabil bleiben muss als der durch Nun modifizierte, der direkt zur Termination führt.

3.3 Die Interaktion des NusB:NusE Heterodimers mit RNA

Die Wechselwirkung zwischen dem NusB:NusE Heterodimer und der *boxA* der *nut* RNA gilt als ein wichtiger Schritt bei der Ausbildung des Antiterminationskomplexes (Greive *et al.*, 2005). NusB ist in der Lage die RNA in Abwesenheit von NusE zu binden, jedoch wird die Affinität durch NusE um etwa den Faktor 10 erhöht (Mason *et al.*, 1992; Nodwell und Greenblatt, 1993; Lüttgen *et al.*, 2002; Greive *et al.*, 2005). Strukturelle Analysen des NusB:NusE Heterodimers wurden lange Zeit durch die geringe Löslichkeit von NusE behindert (Lüttgen *et al.*, 2002) und dies konnte erst durch das Ersetzen einer Ribosomenbindungsschleife von 16 Aminosäuren durch ein einzelnes Serin umgangen werden (Luo *et al.*, 2008). Aktuelle strukturelle Untersuchungen mit dieser NusE Δ genannten Variante im Komplex mit NusB haben gezeigt, dass die RNA-Bindungsfläche primär von NusB gebildet wird, jedoch auch Teile von NusE daran beteiligt sind (Luo *et al.*, 2008). Fluoreszenz-Anisotropie-Messungen mit unterschiedlichsten *nut* und *rrn boxA* Konstrukten liefern einen detaillierten Einblick in die Protein-RNA-Interaktion von NusB und dem NusB:NusE Dimer (Einzelarbeit D). Insbesondere bekannte *boxA* RNA-Mutationen, die *in vivo* oder *in vitro* negative Effekte auf die Wirksamkeit von λ N und/oder Nun haben (Olson *et al.*, 1984; Robledo *et al.*, 1990; Baron und Weisberg, 1992; Patterson *et al.*, 1994), ermöglichten es, wichtige Nukleotide für die Interaktion zu identifizieren. Ein weiterer Punkt war die Untersuchung von NusB101, einer Asp118Asn Mutante. Diese Mutante besitzt die Fähigkeit Mutationen in anderen Nus-Faktoren aufzuheben und somit prozessive Antitermination zu gewährleisten (Ward *et al.*, 1983; Court *et al.*, 1995). Mittels Bestimmung der Kristallstruktur konnte gezeigt werden, dass die Struktur mit der Wildtyp Struktur von NusB:NusE Δ übereinstimmt und eine identische Umorientierung von Rest 118 im Vergleich zum freien NusB zeigt (Einzelarbeit D; Abb. 3-2).

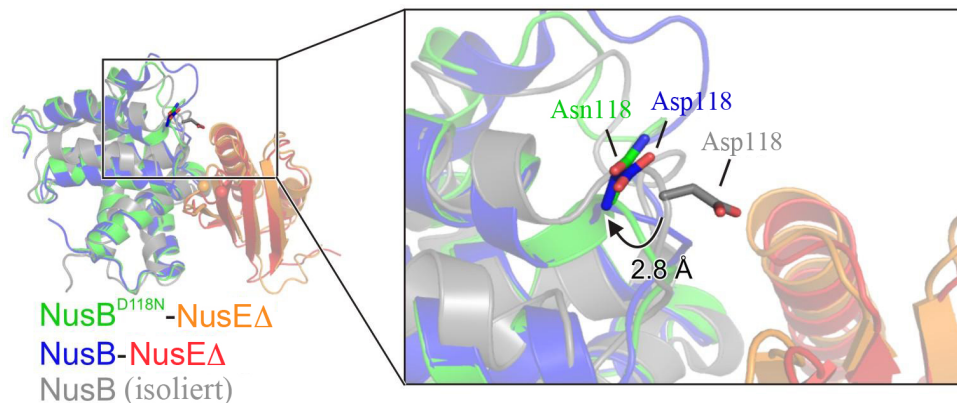


Abbildung 3–2: Effekt der Komplexbildung mit NusEΔ auf NusB

Überlagerung vom NusB:NusEΔ Komplex (Blau und rot; PDB: 3D3B; (Luo *et al.*, 2008)), von NusB (Grau; PDB: 1EY1; (Altieri *et al.*, 2000)) und vom NusB^{D118N}:NusEΔ Komplex (Grün und orange). Reste an Position 118 sind als Sticks dargestellt und in der Vergrößerung hervorgehoben (Kohlenstoff, in der Farbe des entsprechenden Moleküls; Sauerstoff, rot; Stickstoff, blau). Durch Komplexbildung mit NusE wird ein Umklappen des Restes 118 induziert, was einer Bewegung des Cα von 2,8 Å entspricht (Einzelarbeit D, Abbildung 5C).

Die bessere RNA-Bindungsaffinität von NusB101 kann diese anderen Mutationen ausgleichen. Zur genaueren Analyse des Effektes wurden weitere Mutationen, Ala, Arg, Glu, Lys, an dieser Stelle eingeführt (Einzelarbeit D). Strukturell konnten für diese Punktmutationen mittels Circular-Dichroismus-Spektroskopie keine signifikanten Unterschiede zum Wildtyp festgestellt werden. Die Analysen ihrer Bindungsaffinität und ihrer *in vivo* Aktivität zeigten jedoch klare Unterschiede. Zusammenfassend zeigte sich, dass positiv geladene Aminosäuren wie Arg und Lys bevorzugt werden. Dies deutet auf hydrophile Interaktionen mit den Phosphatgruppen der RNA hin, und dass eine gewisse Größe der Aminosäure für die Stabilisierung der RNA an dieser Stelle notwendig ist.

Aus diesen Daten lässt sich zum einen ableiten, dass ein gewisser Grad an Stabilität des NusB:NusE:boxA Komplexes gegeben sein muss, um für prozessive Antitermination zu sorgen. Zum anderen zeigen sie, wie die Feinabstimmung der Rekrutierungsfunktion von NusE durch NusB zur boxA (Luo *et al.*, 2008) im gesamten Antiterminationskomplex erreicht wird (Einzelarbeit D).

3.4 Die einzelnen NusG-Domänen zeigen transiente Interaktionen

Bei der Bestimmung der NMR-Struktur von NusG zeigte sich, dass das gesamte Protein die Tendenz hatte über seine NTD zu aggregieren (Mooney *et al.*, 2009b). Mittels NMR-Relaxationsexperimenten konnte die Rotationskorrelationszeit von NusG bestimmt werden. Hierbei zeigte sich eine bimodale Verteilung mit unterschiedlichem Verhalten für die beiden Domänen (Einzelarbeit E). Die freie NusG-CTD verhält sich wie eine vergleichbar gefaltete SH3-Domäne, wobei es im Gesamtprotein jedoch zu einer deutlichen Erhöhung der Korrelationszeit kommt, da die NTD wie eine Art Anker fungiert. Für die NTD konnte eine deutlich erhöhte Korrelationszeit im Gesamtprotein gezeigt werden, was zusätzlich auf Aggregation in dieser Domäne hindeutet (Einzelarbeit E). Dieser Effekt hängt wahrscheinlich mit diversen hydrophoben Aminosäuren an der Oberfläche der NTD zusammen, die möglicherweise an Protein-Protein-Interaktionen beteiligt sind. Trotz dieser deutlichen Erhöhung weisen die Daten nicht auf eine Oligomerisierung des NusG-Proteins bei Konzentrationen bis 200 μM hin (Aggregation war ab etwa 400 μM deutlich zu beobachten).

Diese hydrophobe Oberfläche um Phe65 auf der NTD wurde in der Literatur schon kontrovers im Zusammenhang mit einem sogenannten *Spring-loaded* Mechanismus diskutiert (Knowlton *et al.*, 2003). Für NusG aus *Aquifex aeolicus* wurde unter bestimmten Kristallisationsbedingungen gezeigt, dass sich ein Dimer, in dem jeweils die NTD mit der CTD des anderen Moleküls wechselwirkt, ausbilden kann (Knowlton *et al.*, 2003). Dieses ist jedoch nur in einem Kristallisationsansatz beobachtbar gewesen (vgl. (Steiner *et al.*, 2002; Knowlton *et al.*, 2003) und es konnte bis *dato* keine Funktion mit diesem postulierten Mechanismus assoziiert werden (Mooney *et al.*, 2009b). Zur Untersuchung dieses möglichen Effektes für *E. coli* NusG wurde die PRE-Methode (vgl. Abschnitt 3.1) angewendet, welche die Möglichkeit bietet, selbst schwache transiente Interaktion zu identifizieren. Für die Untersuchung der einzelnen Domänen wurden die Lys-Reste unselektiv markiert. Bei der Interaktion der isolierten Domänen zeigte sich klar ein Effekt bei beiden Interaktionspartnern (Einzelarbeit E, Abb. 3-3).

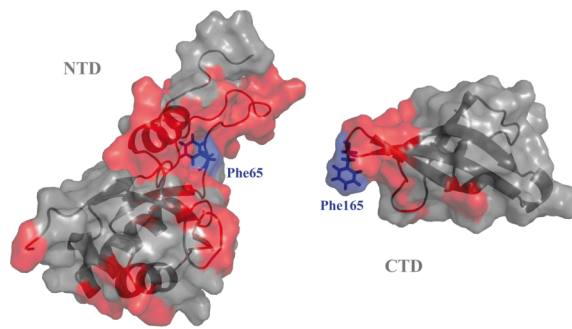


Abbildung 3-3: Effekte des unspezifischen Spin-Labeling der NusG-Domänen

Die beiden zentralen Phenylalanine (Phe65 & Phe165) der transienten Interaktion sind blau hervorgehoben. In rot sind die Reste gekennzeichnet, die eine erhöhte Relaxationsrate aufgrund des Nitroxylradikals ($R_{2,para}(H^N)$) von > 20 Hz bei der Titration mit den einzelnen Domänen aufweisen (Einzelarbeit E, Abbildung 2A)

Aus den Daten der isolierten Domänen lässt sich ableiten, dass die Interaktionsflächen wohl definiert sind, wobei auf NTD Seite aufgrund der bereits erläuterten Aggregationsproblematik der Effekt nicht so lokal begrenzt ist. Es zeigt sich, dass jeweils ein Phenylalanin (Phe65/Phe165), wie es auch für *A. Aeolicus* beschrieben wurde (Knowlton *et al.*, 2003), direkt in diese hydrophobe Interaktion involviert ist. Als nächstes wurde versucht dieses Ergebnis auf das komplette NusG, mittels selektiver Markierung von spezifisch eingeführten Cys-Resten, zu übertragen. Hierbei kann der gleiche Effekt beobachtet werden, es zeigt sich jedoch eine deutliche Konzentrationsabhängigkeit der Interaktion (Einzelarbeit E). Dies belegt, dass die Interaktion nicht in einem Molekül stattfindet und es für NusG nicht wie für sein paraloges Protein RfaH (Belogurov *et al.*, 2007) eine geschlossene Konformation gibt. Andererseits liefert diese Studie neben der bereits erwähnten hydrophoben Oberfläche auf der NTD, für die mittlerweile die Interaktion mit den β' -Klammern der RNAP als gesichert gilt (Belogurov *et al.*, 2009), eindeutige Indizien für die Beteiligung der Reste um Phe165 der CTD an möglichen weiteren Proteininteraktionen innerhalb des Antiterminationskomplexes.

3.5 NusG rekrutiert das NusB:NusE Heterodimer zur RNAP

Neben der bereits angeführten Interaktion mit der RNAP (Belogurov *et al.*, 2009) und der Interaktion mit dem ρ -Faktor (Pasman und von Hippel, 2000), war über die Funktion von NusG innerhalb des Antiterminationskomplexes nichts bekannt. Mittels NMR-Titrations konnte die CTD eindeutig als die ρ -bindende Domäne identifiziert werden (Einzelarbeit F). Zusätzlich konnte zuvor zur Funktion von NusG lediglich gezeigt werden, dass eine Mutation innerhalb der CTD (Ser163Phe) den Effekt anderer Mutationen von NusE bzw. NusA aufheben kann (Sullivan *et al.*, 1992).

Mittels NMR-Titrations ist eindeutig eine Wechselwirkung zwischen der NusG-CTD und NusE Δ in Anwesenheit von NusB nachweisbar (Einzelarbeit F). Aufgrund der Aggregationsproblematik von NusE (vgl. Abschnitt 3.3) ist die Anwesenheit von NusB zwingend erforderlich. Der modulare Aufbau des Komplexes ermöglicht die selektive Markierung der einzelnen Komponenten für NMR-Untersuchungen, wobei für NusE Δ zum Teil auf Deuterierung zur besseren Signalintensität in den Spektren zurückgegriffen werden musste. Auf Basis der unterschiedlichen NMR-Experimente ist die eindeutige Identifikation von mehreren intermolekularen Abstandsbeschränkungen (NOEs) möglich, welche die Interaktion an vier Ankerpunkten fixieren. Damit war es möglich, unter der Annahme fehlender Strukturänderungen der Einzeldomänen, eine eindeutige Struktur durch das Anpassen starrer Körper zu bestimmen (Einzelarbeit F, Abb. 3-4).

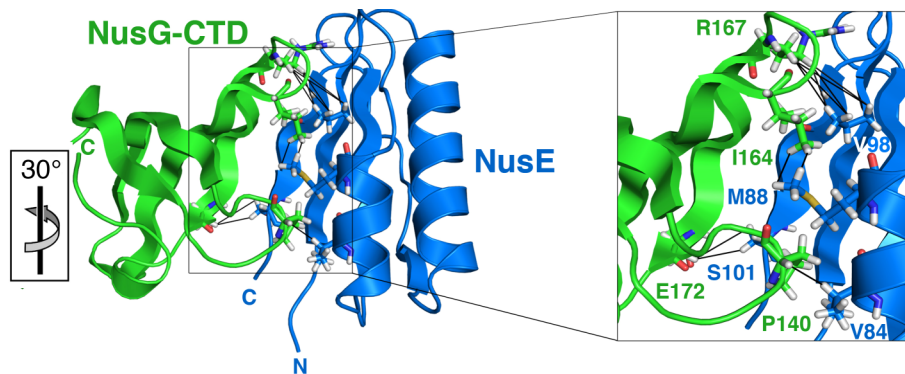


Abbildung 3–4: Experimentelle Basis für die Strukturbestimmung des NusE Δ :NusG-CTD Komplex.

Schwarze Linien zeigen eindeutig bestimmte NOEs zwischen NusG-Arg167-H δ und NusE Δ -Val82-H γ , NusG-Ile164-H δ und NusE Δ -Met72-H ϵ , NusG-Glu172-H β und NusE Δ -Ser85-H β sowie zwischen NusG-Pro140-H α und NusE Δ -Val68-H β . Auf Basis dieser intermolekularen NOEs war eine eindeutige Bestimmung der Orientierung der beiden Proteine zueinander möglich. Zur Übersichtlichkeit wurde auf die Darstellung von NusB verzichtet (Einzelarbeit F, Abbildung 2A).

NusG und NusB binden an den jeweils gegenüberliegenden Seiten von NusE, so dass die Bindung beider Proteine gleichzeitig erfolgen kann, und diese sich nicht gegenseitig behindern (Einzelarbeit F). Die hydrophobe Spitze um Phe165 der NusG-CTD (vgl. 3.5) bindet in eine hydrophobe Vertiefung auf NusE-Seite und wird wahrscheinlich durch weitere Interaktionen, wie z.B. zwischen NusG-Arg167 und NusE-Asp81, stabilisiert. Durch diese Bindung wird das NusB:NusE Heterodimer zur RNAP hin rekrutiert. Die Interaktion ist mit einem K_d von $\sim 50 \mu\text{M}$ (Einzelarbeit F) zwar deutlich schwächer als für die ρ -Interaktion (14 nM (Pasman und von Hippel,

2000)) beschrieben, jedoch ist die NusG:NusE Interaktion in das Netzwerk der verschiedenen Wechselwirkungen zwischen RNAP, NusB, NusE, NusG und *boxA* innerhalb des Antiterminationskomplexes eingebunden, was diesen Unterschied aufheben könnte.

Zusammenfassend stellt sich das aktuelle Bild des Aufbaus des λ Antiterminationskomplexes folgendermaßen dar: NusA ist am Übergang von Initiation zu Elongation durch seine Bindung an die α CTD der RNAP direkt beteiligt. Zusätzlich interagiert es mit unterschiedlichen Domänen mit λ N und der *Spacer*-Region der *nut* RNA. Gleichzeitig bindet λ N an die *boxB* und rekrutiert seinerseits NusA an diese Position. NusB lädt NusE auf die *boxA*. NusG bindet mit seiner CTD an NusE und rekrutiert auf diese Art und Weise bei gleichzeitiger Bindung an die β' -Klammern mittels seiner NTD, wobei es an dieser Stelle den σ -Faktor verdrängt, NusE zur RNAP hin. Dies führt zu einem kompakten und stabilen Antiterminationskomplex (Abb. 3-5).

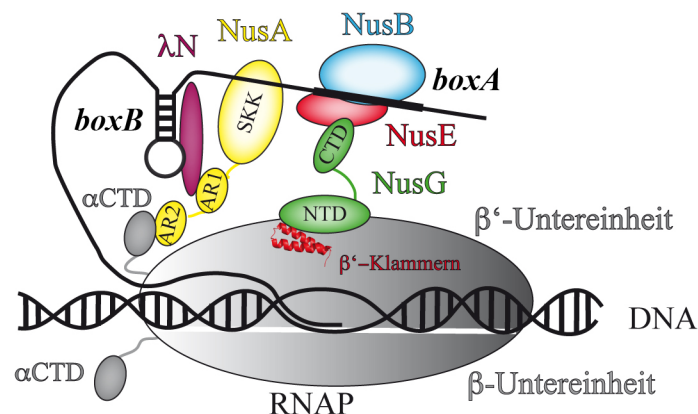


Abbildung 3–5: Aufbau des λ N vermittelten Antiterminationskomplexes.

Aktuelles Bild des molekularen Aufbaus des Antiterminationskomplexes mit seinen Komponenten NusA, NusB, NusE, NusG, RNAP, λ N und *nut* RNA. Details sind im Text aufgeführt.

3.6 Implikationen für die Transkriptions-Translations-Kopplung

Für NusG konnte gezeigt werden, dass es einen direkten Einfluss auf die Translationsgeschwindigkeit hat (Zellars und Squires, 1999). Außerdem konnte die Kopplung von Transkription und Translation in Bakterien bereits bei der Attenuation an den Operons diverser Aminosäuren identifiziert werden (Landick *et al.*, 1996; Yanofsky, 2000). Trotzdem konnte bisher noch keine direkte molekulare Wechselwirkung zwischen beiden Komplexen gezeigt werden.

Neben NusG ist auch NusE als ribosomales Protein S10 direkt an der Translation beteiligt (Squires und Zaporjets, 2000). Nach dem Einbau von NusE/S10 ins Ribosom ist die Bindungsstelle für NusB nicht mehr zugänglich (Luo *et al.*, 2008). Es konnte gezeigt werden, dass eine Überexpression von NusE die Notwendigkeit von NusB zum Aufbau des Antiterminationskomplexes aufheben kann (Luo *et al.*, 2008). Dem gegenüber ist die Bindungsstelle von NusG auf NusE-Seite auch nach Einbau in das Ribosom zugänglich (Einzelarbeit F). Somit könnte die Rolle von NusG die Unterstützung der letzten Schritte des Aufbaus der 30S Untereinheit durch den Einbau von NusE ins Ribosom sein. Durch diese Wechselwirkung wäre NusG das erste Beispiel einer molekularen Verbindung zwischen Transkription und Translation.

4 Abkürzungsverzeichnis

<i>A. aeolicus</i>	<i>Aquifex aeolicus</i>
AR	<i>Acidic repeat</i>
ARM	Arginin-reiches Motiv (<i>arginine-rich motif</i>)
bp	Basenpaare
BH	Brückenhelix
ATP	Adenosintriphosphat
CAP	<i>Catabolite activator protein</i>
CD	Circular Dichroismus
CSP	<i>Chemical shift perturbation</i>
CTD	Carboxy-terminale Domäne
DBS	DNA-Duplex-Bindungsstelle
DNA	Desoxyribonukleinsäure
<i>E. coli</i>	<i>Escherichia Coli</i>
EK	Elongationskomplex
HBS	Hybrid-Bindungsstelle
HSQC	<i>Heteronuclear single quantum coherence</i>
kDa	Kilodalton
KH	K homologe Domäne
MD	Molekular Dynamik
NMR	<i>Nuclear magnetic resonance spectroscopy</i>
NOE	<i>Nuclear overhauser enhancement</i>
nt	Nukleotide
NTD	Amino-terminale Domäne
NTP	Ribonukleosidtriphosphat
Nus A, B, E, G	<i>N utilization substance A, B, E, G</i>
<i>nut</i>	<i>N utilization site</i>
OxyR	<i>Oxydative stress regulator protein</i>
PRE	<i>Paramagnetic relaxation enhancement</i>
RBS	RNA-Bindungsstelle
RNA	Ribonukleinsäure
RNAP	RNA-Polymerase

<i>rrn</i>	Ribosomale RNA
<i>rut</i>	<i>Rho utilization site</i>
S1	S homologe Domäne 1
SD	Shine-Dalgarno Sequenz
<i>T. maritima</i>	<i>Thermatoga maritima</i>
<i>T. thermophilus</i>	<i>Thermus thermophilus</i>
TROSY	<i>transverse relaxation optimized spectroscopy</i>
TS	Trigger-Schleife
UP-Element	<i>upstream promotor element</i>
α CTD	Carboxy-terminale Domäne der α -Untereinheit der RNAP
α NTD	Amino-terminale Domäne der α -Untereinheit der RNAP

5 Literaturverzeichnis

Abbondanzieri EA, Greenleaf WJ, Shaevitz JW, Landick R und Block SM (2005) Direct observation of base-pair stepping by RNA polymerase. *Nature* **438**: 460-465

Aksoy S, Squires CL und Squires C (1984) Evidence for antitermination in *Escherichia coli* rRNA transcription. *J. Bacteriol.* **159**: 260-264

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

Alifano P, Rivellini F, Limauro D, Bruni CB und Carlomagno MS (1991) A consensus motif common to all Rho-dependent prokaryotic transcription terminators. *Cell* **64**: 553-563

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

Artsimovitch I, Landick R (2002) The Transcriptional Regulator RfaH Stimulates RNA Chain Synthesis after Recruitment to Elongation Complexes by the Exposed Nontemplate DNA Strand. *Cell* **109**: 193-203

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

Atsumi S, Little JW (2006) Role of the lytic repressor in prophage induction of phage λ as analyzed by a module-replacement approach. *Proc. Natl. Acad. Sci. U S A* **103**: 4558-4563

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

Bar-Nahum G, Epshtein V, Ruckenstein AE, Rafikov R, Mustaev A und Nudler E (2005) A Ratchet Mechanism of Transcription Elongation and Its Control. *Cell* **120**: 183-193

Baron J, Weisberg RA (1992) Mutations of the phage lambda *nutL* region that prevent the action of Nun, a site-specific transcription termination factor. *J. Bacteriol.* **174**: 1983-9

Batada NN, Westover KD, Bushnell DA, Levitt M und Kornberg RD (2004) Diffusion of nucleoside triphosphates and role of the entry site to the RNA polymerase II active center. *Proc. Natl. Acad. Sci. U S A* **101**: 17361-17364

Belogurov GA, Mooney RA, Svetlov V, Landick R und Artsimovitch I (2009) Functional specialization of transcription elongation factors. *EMBO J.* **28**: 112-122

Belogurov GA, Vassylyeva MN, Svetlov V, Klyuyev S, Grishin NV, Vassylyev DG und Artsimovitch I (2007) Structural Basis for Converting a General

Transcription Factor into an Operon-Specific Virulence Regulator. *Mol. Cell* **26**: 117-129

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

Berg KL, Squires C und Squires CL (1989) Ribosomal RNA operon anti-termination. Function of leader and spacer region *boxB-boxA* sequences and their conservation in diverse micro-organisms. *J. Mol. Biol.* **209**: 345-358

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

Bogden CE, Fass D, Bergman N, Nichols MD und Berger JM (1999) The Structural Basis for Terminator Recognition by the Rho Transcription Termination Factor. *Mol. Cell* **3**: 487-493

Bonin I (2004) Structural and Biochemical Studies of Two Nus Family Proteins: NusB and NusA AR1-N Complex. Doktorarbeit TU München

Bonin I, Mühlberger 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, Nudler E (2008) RNA polymerase: the vehicle of transcription. *Trends Microbiol.* **16**: 126-134

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

Borukhov S, Severinov K (2002) Role of the RNA polymerase sigma subunit in transcription initiation. *Res. Microbiol.* **153**: 557-562

Brennan CA, Dombroski AJ und Platt T (1987) Transcription termination factor rho is an RNA-DNA helicase. *Cell* **48**: 945-952

Brückner F, Cramer P (2008) Structural basis of transcription inhibition by alpha-amanitin and implications for RNA polymerase II translocation. *Nat. Struct. Mol. Biol.* **15**: 811-818

Buck M, Gallegos MT, Studholme DJ, Guo Y und Gralla JD (2000) The bacterial enhancer-dependent sigma⁵⁴ (sigma^N) transcription factor. *J. Bacteriol.* **182**: 4129-4136

Burgess BR, Richardson JP (2001) Transcription Factor Rho Does Not Require a Free End to Act as an RNA-DNA Helicase on an RNA. *J. Biol. Chem.* **276**: 17106-17110

Burgess BR, Richardson JP (2000) RNA Passes through the Hole of the Protein Hexamer in the Complex with the *Escherichia coli* Rho Factor. *J. Biol. Chem.* **276**: 4182-4189

Burgess RR, Anthony L (2001) How sigma docks to RNA polymerase and what sigma does. *Curr. Opin. Microbiol.* **4**: 126-131

- Burns CM, Richardson LV und Richardson JP** (1998) Combinatorial effects of NusA and NusG on transcription elongation and Rho-dependent termination in *Escherichia coli*. *J. Mol. Biol.* **278**: 307-316
- Burova E, Hung SC, Chen J, Court DL, Zhou JG, Mogilnitskiy G und Gottesman ME** (1999) *Escherichia coli* nusG mutations that block transcription termination by coliphage HK022 Nun protein. *Mol. Microbiol.* **31**: 1783-1793
- Burova E, Hung S, Sagitov V, Stitt B und Gottesman M** (1995) *Escherichia coli* NusG protein stimulates transcription elongation rates in vivo and in vitro. *J. Bacteriol.* **177**: 1388-1392
- Cardinale CJ, Washburn RS, Tadigotla VR, Brown LM, Gottesman ME und Nudler E** (2008) Termination factor Rho and its cofactors NusA and NusG silence foreign DNA in *E. coli*. *Science* **320**: 935-938
- Chan CL, Landick R** (1993) Dissection of the his leader pause site by base substitution reveals a multipartite signal that includes a pause RNA hairpin. *J. Mol. Biol.* **233**: 25-42
- Chattopadhyay S, Garcia-Mena J, DeVito J, Wolska K und Das A** (1995) Bipartite function of a small RNA hairpin in transcription antitermination in bacteriophage lambda. *Proc. Natl. Acad. Sci. U S A* **92**: 4061-4065
- 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, Tang C und Iwahara J** (2007) Elucidating transient macromolecular interactions using paramagnetic relaxation enhancement. *Curr. Opin. Struct. Biol.* **17**: 603-616
- Clore GM** (2008) Visualizing lowly-populated regions of the free energy landscape of macromolecular complexes by paramagnetic relaxation enhancement. *Molecular BioSystems* **4**: 1058-1069
- Clore GM, Gronenborn AM** (1998) NMR structure determination of proteins and protein complexes larger than 20 kDa. *Curr. Opin. Chem. Biol.* **2**: 564-570
- 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, Liveris D, Squires C, Schwartz I und Squires C** (1995) rRNA operon multiplicity in *Escherichia coli* and the physiological implications of rrn inactivation. *J. Bacteriol.* **177**: 4152-4156
- 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
- Court DL, Patterson TA, Baker T, Costantino N, Mao X und Friedman DI** (1995) Structural and functional analyses of the transcription-translation proteins NusB and NusE. *J. Bacteriol.* **177**: 2589-2591
- Cramer P, Bushnell DA und Kornberg RD** (2001) Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. *Science* **292**: 1863-1876
- Culver GM, Noller HF** (1999) Efficient reconstitution of functional *Escherichia coli* 30S ribosomal subunits from a complete set of recombinant small subunit ribosomal proteins. *RNA* **5**: 832-843
- Das A** (1992) How the phage lambda N gene product suppresses transcription termination: communication of RNA polymerase with regulatory proteins mediated by signals in nascent RNA. *J. Bacteriol.* **174**: 6711-6716
- Das A, 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
- Das R, Loss S, Li J, Waugh DS, Tarasov S, Wingfield PT, Byrd RA und Altieri AS** (2008) Structural Biophysics of the NusB:NusE Antitermination Complex. *J. Mol. Biol.* **376**: 705-720
- Datta K, von Hippel PH** (2008) Direct Spectroscopic Study of Reconstituted Transcription Complexes Reveals That Intrinsic Termination Is Driven Primarily by Thermodynamic Destabilization of the Nucleic Acid Framework. *J. Biol. Chem.* **283**: 3537-3549
- Dodd IB, Shearwin KE und Egan JB** (2005) Revisited gene regulation in bacteriophage lambda. *Curr. Opin. Genet. Dev.* **15**: 145-152
- Donahue J, Turnbough C, Jr** (1994) Nucleotide-specific transcriptional pausing in the pyrBI leader region of *Escherichia coli* K-12. *J. Biol. Chem.* **269**: 18185-18191
- Ebright RH** (2000) RNA Polymerase: Structural Similarities Between Bacterial RNA Polymerase and Eukaryotic RNA Polymerase II. *J. Mol. Biol.* **304**: 687-698
- Eisenmann A, Schwarz S, Prasad S, Schweimer K und Rösch P** (2005) The *E. coli* NusA carboxy-terminal domains are structurally similar and show specific RNAP- and λ N interaction. *Protein Sci.* **14**: 2018-2029
- Ellwood M, Nomura M** (1982) Chromosomal locations of the genes for rRNA in *Escherichia coli* K-12. *J. Bacteriol.* **149**: 458-468
- Epshtein V, Cardinale CJ, Ruckenstein AE, Borukhov S und Nudler E** (2007) An allosteric path to transcription termination. *Mol. Cell* **28**: 991-1001
- Epshtein V, Nudler E** (2003) Cooperation between RNA polymerase molecules in transcription elongation. *Science* **300**: 801-805
- 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
- Faber C, Schärpf M, Becker T, Sticht H und Rösch P** (2001) The structure of the coliphage HK022 N protein-lambda-phage *boxB* RNA complex. Implications for the mechanism of transcription termination. *J. Biol. Chem.* **276**: 32064-32070
- Fiaux J, Bertelsen EB, Horwich AL und Wüthrich K** (2002) NMR analysis of a 900K GroEL-GroES complex. *Nature* **418**: 207-211
- Fish RN, Kane CM** (2002) Promoting elongation with transcript cleavage stimulatory factors. *Biochim. Biophys. Acta* **1577**: 287-307
- Frank J, Agrawal RK** (2000) A ratchet-like inter-subunit reorganization of the ribosome during translocation. *Nature* **406**: 318-322
- Frank J, Gao H, Sengupta J, Gao N und Taylor DJ** (2007) The process of mRNA-tRNA translocation. *Proc. Natl. Acad. Sci. U S A* **104**: 19671-19678
- Franklin NC** (1993) Clustered Arginine Residues of Bacteriophage λ N Protein are Essential to Antitermination of Transcription, but Their Locale Cannot Compensate for *boxB* Loop Defects. *J. Mol. Biol.*, **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, Court DL** (1995) Transcription antitermination: the lambda paradigm updated. *Mol. Microbiol.* **18**: 191-200
- 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
- Galburt EA, Grill SW, Wiedmann A, Lubkowska L, Choy J, Nogales E, Kashlev M und Bustamante C** (2007) Backtracking determines the force sensitivity of RNAP II in a factor-dependent manner. *Nature* **446**: 820-823
- 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
- Goldman SR, Ebright RH und Nickels BE** (2009) Direct Detection of Abortive RNA Transcripts *in vivo*. *Science* **324**: 927-928
- 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, 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
- 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 antitermination 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
- Greive SJ, Weitzel SE, Goodarzi JP, Main LJ, Pasman Z und von Hippel PH** (2008) Monitoring RNA transcription in real time by using surface plasmon resonance. *Proc. Natl. Acad. Sci. U S A* **105**: 3315-3320
- 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
- Hartzog GA, Wada T, Handa H und Winston F** (1998) Evidence that Spt4, Spt5, and Spt6 control transcription elongation by RNA polymerase II in *Saccharomyces cerevisiae*. *Genes Dev.* **12**: 357-369
- Hasan N, Szybalski W** (1986) Effect of the promoter structure on the *nutL* transcription antitermination function. *Gene* **50**: 97-100
- Helmann JD** (1999) Anti-sigma factors. *Curr. Opin. Microbiol.* **2**: 135-141
- Helmann JD, Chamberlin MJ** (1988) Structure and function of bacterial sigma factors. *Annu. Rev. Biochem.* **57**: 839-872
- 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
- Hsu LM, Vo NV, Kane CM und Chamberlin MJ** (2003) *In vitro* Studies of Transcript Initiation by *Escherichia coli* RNA Polymerase. 1. RNA Chain Initiation, Abortive Initiation, and Promoter Escape at Three Bacteriophage Promoters. *Biochemistry* **42**: 3777-3786
- 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

- Jucker FM, Heus HA, Yip PF, Moors EHM und Pardi A** (1996) A Network of Heterogeneous Hydrogen Bonds in GNRA Tetraloops. *J. Mol. Biol.*, **264**: 968-980
- Kaiser AD** (1957) Mutations in a temperate bacteriophage affecting its ability to lysogenize *Escherichia coli*. *Virology* **3**: 42-61
- Kettenberger H, Armache K und Cramer P** (2004) Complete RNA Polymerase II Elongation Complex Structure and Its Interactions with NTP and TFIIIS. *Mol. Cell* **16**: 955-965
- Kim D, Patel SS** (2001) The Kinetic Pathway of RNA Binding to the *Escherichia coli* Transcription Termination Factor Rho. *J. Biol. Chem.* **276**: 13902-13910
- Kireeva ML, Hancock B, Cremona GH, Walter W, Studitsky VM und Kashlev M** (2005) Nature of the Nucleosomal Barrier to RNA Polymerase II. *Mol. Cell* **18**: 97-108
- Kireeva ML, Kashlev M** (2009) Mechanism of sequence-specific pausing of bacterial RNA polymerase. *Proc. Natl. Acad. Sci. U S A* **106**: 8900-8905
- Knowlton JR, Bubunenko M, Andrykovitch M, Guo W, Routzahn KM, Waugh DS, Court DL und Ji X** (2003) A Spring-Loaded State of NusG in Its Functional Cycle Is Suggested by X-ray Crystallography and Supported by Site-Directed Mutants. *Biochemistry* **42**: 2275-2281
- Komissarova N, Kashlev M** (1997a) RNA polymerase switches between inactivated and activated states by translocating back and forth along the DNA and the RNA. *J. Biol. Chem.* **272**: 15329-15338
- Komissarova N, Kashlev M** (1997b) Transcriptional arrest: *Escherichia coli* RNA polymerase translocates backward, leaving the 3' end of the RNA intact and extruded. *Proc. Natl. Acad. Sci. U S A* **94**: 1755-1760
- Korzheva N, Mustaev A, Kozlov M, Malhotra A, Nikiforov V, Goldfarb A und Darst SA** (2000) A structural model of transcription elongation. *Science* **289**: 619-625
- Koulich D, Orlova M, Malhotra A, Sali A, Darst SA und Borukhov S** (1997) Domain Organization of *Escherichia coli* Transcript Cleavage Factors GreA and GreB. *J. Biol. Chem.* **272**: 7201-7210
- Kuznedelov K, Korzheva N, Mustaev A und Severinov K** (2002) Structure-based analysis of RNA polymerase function: the largest subunit's rudder contributes critically to elongation complex stability and is not involved in the maintenance of RNA-DNA hybrid length. *EMBO J.* **21**: 1369-1378
- Kyrpides NC, Woese CR und Ouzounis CA** (1996) KOW: a novel motif linking a bacterial transcription factor with ribosomal proteins. *Trends Biochem. Sci.* **21**: 425-426
- Kyrpides NC, Ouzounis CA** (1999) Transcription in Archaea. *Proc. Natl. Acad. Sci. U S A* **96**: 8545-8550
- Landick R** (2006) The regulatory roles and mechanism of transcriptional pausing. *Biochem. Soc. Trans.* **34**: 1062-1066
- Landick R, Turnbough CL, Jr und Yanofsky C** (1996) Transcription attenuation. In FC Neidhardt, R Curtiss, J Ingraham, ECC Lin, KB Low, eds, *Escherichia coli* and

Salmonella: Cellular and Molecular Biology, Ed 2nd Vol 1. *Am. Soc. Microbiol.*, Washington, DC, pp 1263-1286

Landick R, Yanofsky C (1987) Isolation and structural analysis of the *Escherichia coli* trp leader paused transcription complex. *J. Mol. Biol.* **196**: 363-377

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

Larson MH, Greenleaf WJ, Landick R und Block SM (2008) Applied Force Reveals Mechanistic and Energetic Details of Transcription Termination. *Cell* **132**: 971-982

Lazinski D, Grzadzielska E und Das A (1989) Sequence-specific recognition of RNA hairpins by bacteriophage antiterminators requires a conserved arginine-rich motif. *Cell*, **59**: 207-218

Lederberg EM (1951) Lysogenicity in *E. coli* K.12. *Genetics* **36**: 560

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

Lesnik EA, Sampath R, Levene HB, Henderson TJ, McNeil JA und Ecker DJ (2001) Prediction of rho-independent transcriptional terminators in *Escherichia coli*. *Nucl. Acids Res.* **29**: 3583-3594

Li J, Horwitz R, McCracken S und Greenblatt J (1992) NusG, a new *Escherichia coli* elongation factor involved in transcriptional antitermination by the N protein of phage lambda. *J. Biol. Chem.* **267**: 6012-9

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

Lin YC, Choi WS und Gralla JD (2005) TFIIF XPB mutants suggest a unified bacterial-like mechanism for promoter opening but not escape. *Nat. Struct. Mol. Biol.* **12**: 603-607

Lisser S, Margalit H (1993) Compilation of *E. coli* mRNA promoter sequences. *Nucl. Acids Res.* **21**: 1507-1516

Little JW (2005) Threshold effects in gene regulation: When some is not enough. *Proc. Natl. Acad. Sci. U. S. A.* **102**: 5310-5311

Liu K, Hanna MM (1995) NusA contacts nascent RNA in *Escherichia coli* transcription complexes. *J. Mol. Biol.* **247**: 547-558

Luo X, Hsiao HH, Bubunenko M, Weber G, Court DL, Gottesman ME, Urlaub H und Wahl MC (2008) Structural and functional analysis of the *E. coli* NusB-S10 transcription antitermination complex. *Mol. Cell* **32**: 791-802

Lüttgen H, Robelek R, Mühlberger R, Diercks T, Schuster SC, Kohler P, Kessler H, Bacher A und Richter G (2002) Transcriptional regulation by antitermination.

Interaction of RNA with NusB protein and NusB/NusE protein complex of *Escherichia coli*. *J. Mol. Biol.* **316**: 875-885

Macdonald LE, Zhou Y und McAllister WT (1993) Termination and slippage by bacteriophage T7 RNA polymerase. *J. Mol. Biol.* **232**: 1030-1047

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

Martin FH, Tinoco I, Jr. (1980) DNA-RNA hybrid duplexes containing oligo (dA:rU) sequences are exceptionally unstable and may facilitate termination of transcription. *Nucl. Acids Res.* **8**: 2295-2300

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

Mathew R, Chatterji D (2006) The evolving story of the omega subunit of bacterial RNA polymerase. *Trends Microbiol.* **14**: 450-455

Maurizi M (1987) Degradation *in vitro* of bacteriophage lambda N protein by Lon protease from *Escherichia coli*. *J. Biol. Chem.* **262**: 2696-2703

Mejia YX, Mao H, Forde NR und Bustamante C (2008) Thermal Probing of *E. coli* RNA Polymerase Off-Pathway Mechanisms. *J. Mol. Biol.* **382**: 628-637

Metzger W, Schickor P, Meier T, Werel W und Heumann H (1993) Nucleation of RNA Chain Formation by *Escherichia coli* DNA-dependent RNA Polymerase. *J. Mol. Biol.*, **232**: 35-49

Minakhin L, Bhagat S, Brunning A, Campbel EA, Darst SA, Ebright RH und Severinov K (2001) Bacterial RNA polymerase subunit ω and eukaryotic RNA polymerase subunit RPB6 are sequence, structural, and functional homologs and promote RNA polymerase assembly. *Proc. Natl. Acad. Sci. U S A* **98**: 892-897

Mittermaier A, Kay LE (2006) New Tools Provide New Insights in NMR Studies of Protein Dynamics. *Science* **312**: 224-228

Mizushima S, Nomura M (1970) Assembly mapping of 30S ribosomal proteins from *E. coli*. *Nature* **226**: 1214

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

- Mooney RA, Davis SE, Peters JM, Rowland JL, Ansari AZ und Landick R** (2009a) Regulator Trafficking on Bacterial Transcription Units *in vivo*. *Mol. Cell* **33**: 97-108
- Mooney RA, Schweimer K, Rösch P, Gottesman ME und Landick R** (2009b) Two Structurally Independent Domains of *E. coli* NusG Create Regulatory Plasticity via Distinct Interactions with RNA Polymerase and Regulators. *J. Mol. Biol.* **391**: 341-358
- Morgan EA** (1986) Antitermination mechanisms in rRNA operons of *Escherichia coli*. *J. Bacteriol.* **168**: 1-5
- Mühlberger R, Robelek R, Eisenreich W, Ettenhuber C, Sinner EK, Kessler H, Bacher A und Richter G** (2003) RNA DNA Discrimination by the Antitermination Protein NusB. *J. Mol. Biol.*, **327**: 973-983
- 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 KS, Darst SA** (2003) Bacterial RNA polymerases: the whole story. *Curr. Opin. Struct. Biol.* **13**: 31-39
- Mustaev A, Kozlov M, Markovtsov V, Zaychikov E, Denissova L und Goldfarb A** (1997) Modular organization of the catalytic center of RNA polymerase. *Proc. Natl. Acad. Sci. U S A* **94**: 6641-6645
- 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
- Nodwell JR, Greenblatt J** (1993) Recognition of *boxA* antiterminator RNA by the *E. coli* antitermination factors NusB and ribosomal protein S10. *Cell* **72**: 261-268
- Noller HF, Nomura M** (1996) Ribosomes. In FC Neidhardt, R Curtiss, J Ingraham, ECC Lin, KB Low, eds, *Escherichia Coli* and *Salmonella*: Cellular and Molecular Biology, Ed 2nd Vol 1. *Am. Soc. Microbiol.*, Washington, D.C., USA, pp 167-186
- Nudler E** (2009) RNA Polymerase Active Center: The Molecular Engine of Transcription. *Annu. Rev. Biochem.* **78**: 335-361
- Nudler E** (1999) Transcription elongation: structural basis and mechanisms. *J. Mol. Biol.* **288**: 1-12
- Nudler E, Avetisova E, Markovtsov V und Goldfarb A** (1996) Transcription processivity: protein-DNA interactions holding together the elongation complex. *Science* **273**: 211-217
- 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

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

Nudler E, Mustaev A, Goldfarb A und Lukhtanov E (1997) The RNA-DNA Hybrid Maintains the Register of Transcription by Preventing Backtracking of RNA Polymerase. *Cell* **89**: 33-41

Oberto J, Weisberg RA und Gottesman ME (1989) Structure and function of the *nun* gene and the immunity region of the lambdoid phage HK022. *J. Mol. Biol.* **207**: 675-693

Olson ER, Tomich CS und Friedman DI (1984) The nusA recognition site. Alteration in its sequence or position relative to upstream translation interferes with the action of the N antitermination function of phage lambda. *J. Mol. Biol.* **180**: 1053-63

Pagadala Santhanam S (2008) Structural insights into RNA binding by NusA and interaction studies of Nun with *E. coli* Nus factors. Doktorarbeit Universität Bayreuth

Paget MS, Helmann JD (2003) The sigma⁷⁰ family of sigma factors. *Genome Biol.* **4**: 203

Pasman Z, von Hippel PH (2000) Regulation of Rho-Dependent Transcription Termination by NusG Is Specific to the *Escherichia coli* Elongation Complex. *Biochemistry* **39**: 5573-5585

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

Pervushin K, Riek R, Wider G und Wüthrich K (1997) Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proc. Natl. Acad. Sci. U S A* **94**: 12366-12371

Prasch S (2008) Die Regulation der prokaryontischen Transkription auf molekularer Ebene. Doktorarbeit Universität Bayreuth

Prasch S, Jurk M, Washburn RS, Gottesman ME, Wöhrl BM und Rösch P (2009) RNA-binding specificity of *E. coli* NusA. *Nucl. Acids Res.* *in press*

Prasch S, Schwarz S, Eisenmann A, Wöhrl BM, Schweimer K und Rösch P (2006) Interaction of the intrinsically unstructured phage lambda N Protein with *E. coli* NusA. *Biochemistry* **45**: 4542-4549

Ptashne M (2006) Lambda's switch: lessons from a module swap. *Curr. Biol.* **16**: R459-462

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

Reay P, Yamasaki K, Terada T, Kuramitsu S, Shirouzu M und Yokoyama S (2004) Structural and sequence comparisons arising from the solution structure of the transcription elongation factor NusG from *Thermus thermophilus*. *Proteins* **56**: 40-51

Reeder TC, Hawley DK (1996) Promoter Proximal Sequences Modulate RNA Polymerase II Elongation by a Novel Mechanism. *Cell* **87**: 767-777

- Rees WA, Weitzel SE, Das A und von Hippel PH** (1997) Regulation of the elongation-termination decision at intrinsic terminators by antitermination protein N of phage lambda. *J. Mol. Biol.* **273**: 797-813
- 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
- Reppas NB, Wade JT, Church G und Struhl K** (2006) The Transition between Transcriptional Initiation and Elongation in *E. coli* Is Highly Variable and Often Rate Limiting. *Mol. Cell* **24**: 747-757
- Richardson JP** (2002) Rho-dependent termination and ATPases in transcript termination. *Biochim. Biophys. Acta* **1577**: 251-260
- Robert J, Sloan SB, Weisberg RA, Gottesman ME, Robledo R und Harbrecht D** (1987) The remarkable specificity of a new transcription termination factor suggests that the mechanisms of termination and antitermination are similar. *Cell* **51**: 483-492
- Roberts JW** (1969) Termination factor for RNA synthesis. *Nature* **224**: 1168-1174
- Roberts JW, Shankar S und Filter JJ** (2008) RNA polymerase elongation factors. *Annu. Rev. Microbiol.* **62**: 211-233
- Roberts JW, Yarnell W, Bartlett E, Guo J, Marr M, Ko DC, Sun H und Roberts CW** (1998) Antitermination by bacteriophage lambda Q protein. *Cold Spring Harb. Symp. Quant. Biol.* **63**: 319-325
- Robledo R, Atkinson BL und Gottesman ME** (1991) *Escherichia coli* mutations that block transcription termination by phage HK022 Nun protein. *J. Mol. Biol.* **220**: 613-619
- Robledo R, Gottesman ME und Weisberg RA** (1990) Lambda *nutR* mutations convert HK022 Nun protein from a transcription termination factor to a suppressor of termination. *J. Mol. Biol.* **212**: 635-43
- 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
- Sakata-Sogawa K, Shimamoto N** (2004) RNA polymerase can track a DNA groove during promoter search. *Proc. Natl. Acad. Sci. U S A* **101**: 14731-14735
- Salzmann M, Pervushin K, Wider G, Senn H und Wüthrich K** (1998) TROSY in triple-resonance experiments: new perspectives for sequential NMR assignment of large proteins. *Proc. Natl. Acad. Sci. U S A* **95**: 13585-13590
- Santangelo TJ, Roberts JW** (2004) Forward translocation is the natural pathway of RNA release at an intrinsic terminator. *Mol. Cell* **14**: 117-126
- Schärpf M, Sticht H, Schweimer K, Boehm M, Hoffmann S und Rösch P** (2000) Antitermination in bacteriophage lambda. The structure of the N36 peptide-*boxB* RNA complex. *Eur. J. Biochem.* **267**: 2397-2408
- Schluenzen F, Tocilj A, Zarivach R, Harms J, Gluehmann M, Janell D, Bashan A, Bartels H, Agmon I, Franceschi F und Yonath A** (2000) Structure of functionally activated small ribosomal subunit at 3.3 angstroms resolution. *Cell* **102**: 615-623

- Schmidt MC, Chamberlin MJ** (1987) NusA protein of *Escherichia coli* is an efficient transcription termination factor for certain terminator sites. *J. Mol. Biol.* **195**: 809-818
- 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
- Schroeder LA, Choi AJ und deHaseth P** (2007) The -11A of promoter DNA and two conserved amino acids in the melting region of σ^{70} both directly affect the rate limiting step in formation of the stable RNA polymerase-promoter complex, but they do not necessarily interact. *Nucl. Acids Res.* **35**: 4141-4153
- Schroeder LA, Karpen ME und deHaseth P** (2008) Threonine 429 of *E. coli* σ^{70} is a key participant in promoter DNA melting by RNA polymerase. *J. Mol. Biol.* **376**: 153-165
- Schuwirth BS, Borovinskaya MA, Hau CW, Zhang W, Vila-Sanjurjo A, Holton JM und Cate JHD** (2005) Structures of the Bacterial Ribosome at 3.5 Å Resolution. *Science* **310**: 827-834
- Severinov K** (2000) RNA polymerase structure-function: insights into points of transcriptional regulation. *Curr. Opin. Microbiol.* **3**: 118-125
- Sevostyanova A, Svetlov V, Vassylyev DG und Artsimovitch I** (2008) The elongation factor RfaH and the initiation factor sigma bind to the same site on the transcription elongation complex. *Proc. Nat. Acad. Sci. U S A* **105**: 865-870
- 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
- Skordalakes E, Berger JM** (2006) Structural Insights into RNA-Dependent Ring Closure and ATPase Activation by the Rho Termination Factor. *Cell* **127**: 553-564
- Sosunov V, Zorov S, Sosunova E, Nikolaev A, Zakeyeva I, Bass I, Goldfarb A, Nikiforov V, Severinov K und Mustaev A** (2005) The involvement of the aspartate triad of the active center in all catalytic activities of multisubunit RNA polymerase. *Nucl. Acids Res.* **33**: 4202-4211
- Squires CL, Greenblatt J, Li J, Condon C und Squires CL** (1993) Ribosomal RNA antitermination in vitro: requirement for Nus factors and one or more unidentified cellular components. *Proc. Natl. Acad. Sci. U S A* **90**: 970-974
- Squires CL, Zaporjets D** (2000) Proteins shared by the transcription and translation machines. *Annu. Rev. Microbiol.* **54**: 775-798
- Steiner T, Kaiser JT, Marinkovic S, Huber R und Wahl MC** (2002) Crystal structures of transcription factor NusG in light of its nucleic acid- and protein-binding activities. *EMBO J.* **21**: 4641-4653
- Stuart AC, Gottesman ME und Palmer AG, 3rd** (2003) The N-terminus is unstructured, but not dynamically disordered, in the complex between HK022 Nun protein and lambda-phage *boxB* RNA hairpin. *FEBS Lett.* **553**: 95-8

- 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 anti-termination in phage lambda. *Biochemistry* **36**: 12722-12732
- Sullivan SL, Gottesman ME** (1992) Requirement for *E. coli* NusG protein in factor-dependent transcription termination. *Cell* **68**: 989-994
- Sullivan SL, Ward DF und Gottesman ME** (1992) Effect of *Escherichia coli* nusG function on lambda N-mediated transcription antitermination. *J. Bacteriol.* **174**: 1339-1344
- 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
- Torres M, Balada JM, Zellars M, Squires C und Squires CL** (2004) *In vivo* effect of NusB and NusG on rRNA transcription antitermination. *J. Bacteriol.* **186**: 1304-1310
- Torres M, Condon C, Balada JM, Squires C und Squires CL** (2001) Ribosomal protein S4 is a transcription factor with properties remarkably similar to NusA, a protein involved in both non-ribosomal and ribosomal RNA antitermination. *Embo J.* **20**: 3811-20
- 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
- Touloukhonov I, Zhang J, Palangat M und Landick R** (2007) A Central Role of the RNA Polymerase Trigger Loop in Active-Site Rearrangement during Transcriptional Pausing. *Mol. Cell* **27**: 406-419
- Traviglia 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
- Van Gilst MR, Rees WA, Das A und von Hippel PH** (1997) Complexes of N Antitermination Protein of Phage λ with Specific and Nonspecific RNA Target Sites on the Nascent Transcript. *Biochemistry* **36**: 1514-1524
- 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
- Vassylyev DG, Sekine S, Laptenko O, Lee J, Vassylyeva MN, Borukhov S und Yokoyama S** (2002) Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 Å resolution. *Nature* **417**: 712-719
- Vassylyev DG, Vassylyeva MN, Perederina A, Tahirov TH und Artsimovitch I** (2007a) Structural basis for transcription elongation by bacterial RNA polymerase. *Nature* **448**: 157-162

- Vassylyev DG, Vassylyeva MN, Zhang J, Palangat M, Artsimovitch I und Landick R** (2007b) Structural basis for substrate loading in bacterial RNA polymerase. *Nature* **448**: 163-168
- Vassylyeva MN, Svetlov V, Dearborn AD, Klyuyev S, Artsimovitch I und Vassylyev DG** (2007) The carboxy-terminal coiled-coil of the RNA polymerase β' -subunit is the main binding site for Gre factors. *EMBO Rep.* **8**: 1038-1043
- 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, Pasmán Z** (2002) Reaction pathways in transcript elongation. *Biophys. Chem.* **101-102**: 401-423
- Wang D, Bushnell DA, Westover KD, Kaplan CD und Kornberg RD** (2006) Structural basis of transcription: role of the trigger loop in substrate specificity and catalysis. *Cell* **127**: 941-954
- Wang M, Schnitzer MJ, Yin H, Landick R, Gelles J und Block SM** (1998) Force and velocity measured for single molecules of RNA polymerase. *Science* **282**: 902-907
- Ward DF, DeLong A und Gottesman ME** (1983) *Escherichia coli* nusB mutations that suppress nusA1 exhibit lambda N specificity. *J. Mol. Biol.* **168**: 73-85
- Watnick RS, Gottesman ME** (1999) Binding of transcription termination protein nun to nascent RNA and template DNA. *Science* **286**: 2337-2339
- Watnick RS, Gottesman ME** (1998) *Escherichia coli* NusA is required for efficient RNA binding by phage HK022 nun protein. *Proc. Natl. Acad. Sci. U S A* **95**: 1546-1551
- Watnick RS, Herring SC, Palmer AG, 3rd und Gottesman ME** (2000) The carboxyl terminus of phage HK022 Nun includes a novel zinc-binding motif and a tryptophan required for transcription termination. *Genes Dev.* **14**: 731-739
- Weisberg RA** (2008) Transcription by Moonlight: Structural Basis of an Extraribosomal Activity of Ribosomal Protein S10. *Mol. Cell* **32**: 747-748
- Weisberg RA, Gottesman ME** (1999) Processive antitermination. *J. Bacteriol.* **181**: 359-367
- Westover KD, Bushnell DA und Kornberg RD** (2004) Structural Basis of Transcription: Nucleotide Selection by Rotation in the RNA Polymerase II Active Center. *Cell* **119**: 481-489
- 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
- Wilson HR, Zhou JG, Yu D und Court DL** (2004) Translation repression by an RNA polymerase elongation complex. *Mol. Microbiol.* **53**: 821-828
- Wimberly BT, Brodersen DE, Clemons WMJ, Morgan-Warren RJ, Carter AP, Vornheim C, Hartsch T und Ramakrishnan V** (2000) Structure of the 30S ribosomal subunit. *Nature* **407**: 327-339
- Winkler ME, Yanofsky C** (1981) Pausing of RNA polymerase during *in vitro* transcription of the tryptophan operon leader region. *Biochemistry* **20**: 3738-3744

- 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
- Yakhnin AV, Yakhnin H und Babitzke P** (2006) RNA Polymerase Pausing Regulates Translation Initiation by Providing Additional Time for TRAP-RNA Interaction. *Mol. Cell* **24**: 547-557
- Yanofsky C, Konan KV und Sarsero JP** (1996) Some novel transcription attenuation mechanisms used by bacteria. *Biochimie* **78**: 1017-1024
- Yanofsky C** (2000) Transcription Attenuation: Once Viewed as a Novel Regulatory Strategy. *J. Bacteriol.* **182**: 1-8
- Yarnell WS, Roberts JW** (1999) Mechanism of intrinsic transcription termination and antitermination. *Science* **284**: 611-5
- Yu X, Horiguchi T, Shigesada K und Egelman EH** (2000) Three-dimensional reconstruction of transcription termination factor rho: orientation of the N-terminal domain and visualization of an RNA-binding site. *J. Mol. Biol.* **299**: 1279-1287
- Zellars M, Squires CL** (1999) Antiterminator-dependent modulation of transcription elongation rates by NusB and NusG. *Mol. Microbiol.* **32**: 1296-304
- 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
- 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
- Zuiderweg ER** (2002) Mapping protein-protein interactions in solution by NMR spectroscopy. *Biochemistry* **41**: 1-7

6 Publikationsliste

6.1 Einzelarbeit A

Björn M. Burmann, Augusto Uc-Mass, Kristian Schweimer, Max E. Gottesman und Paul Rösch (2008): The Y39A Mutation of HK022 Nun Disrupts a *boxB* Interaction but Preserves Termination Activity. *Biochemistry* **47**, 7335-7341

Von mir wurden die Fluoreszenz- und die NMR-Experimente sowie die Molekular-Dynamik (MD) Berechnungen durchgeführt. Die NMR-Experimente wurden mit Hilfe von Kristian Schweimer gemessen und analysiert. Die *in vivo* Experimente wurden von Augusto Uc-Mass im Labor von Max E. Gottesman durchgeführt. Das Manuskript wurde von Paul Rösch, Kristian Schweimer, Max E. Gottesman und mir verfasst.

6.2 Einzelarbeit B

Irena Matečko, Björn M. Burmann, Kristian Schweimer, Hubert Kalbacher, Jürgen Einsiedel, Peter Gmeiner und Paul Rösch (2008): Structural Characterisation of the *E. coli* Heat Stable Enterotoxin STh. *The Open Spectroscopy Journal* **2**, 34-39

Die Strukturbestimmung mittels NMR wurde von Irena Matečko unter der Mithilfe von Kristian Schweimer durchgeführt. Das STh wurde von Hubert Kalbacher, Jürgen Einsiedel und Peter Gmeiner chemisch synthetisiert. MD-Berechnungen zu STh und den verwandten Peptiden Guanylin und Uroguanylin wurden von mir durchgeführt. Das Manuskript wurde von Paul Rösch, Irena Matečko, Hubert Kalbacher und mir verfasst.

6.3 Einzelarbeit C

Claudia Knake, Björn M. Burmann, Kristian Schweimer, Irena Matečko und Paul Rösch (2009): The Membrane Proximal Extracellular Domain of Human hGC-B Folds Independently. *Journal of Biomolecular Structure & Dynamics* **26**, 465-472

Die Expression und Reinigung von hGC-B wurde von Claudia Knake unter der Anleitung von Irena Matečko durchgeführt. Die Homologie-Modellierung wurde von Irena Matečko und die MD Berechnungen zur Domänenstabilität wurden von mir durchgeführt. Die NMR-Experimente wurden von Claudia Knake mit Hilfe von Irena Matečko und Kristian Schweimer durchgeführt und analysiert. Das Manuskript wurde von Paul Rösch, Claudia Knake, Irena Matečko, Kristian Schweimer und mir verfasst.

6.4 Einzelarbeit D

Björn M. Burmann, Xiao Luo, Paul Rösch, Markus C. Wahl und Max E. Gottesman (2010): Fine tuning of the *E. coli* NusB:NusE complex affinity to RNA is required for processive antitermination. *Nucleic Acids Research* **38**, 314 – 326

Die aufgeführten Experimente wurden von mir geplant und zum Großteil selbst durchgeführt. Die Fluoreszenzanisotropie- und Circular-Dichroismus-Messungen wurden von mir durchgeführt. Die *in vivo* Experimente wurden von mir unter der Anleitung von Max E. Gottesman im Rahmen eines Forschungsaufenthaltes an der Columbia Universität durchgeführt. Xiao Luo hat den NusB101:NusE Komplex kristallisiert. Das Manuskript wurde von Paul Rösch, Max E. Gottesman, Markus C. Wahl und mir verfasst.

6.5 Einzelarbeit E

Björn M. Burmann, Ulrich Scheckenhofer, Kristian Schweimer und Paul Rösch (2010): Domain interactions of the transcription:translation coupling factor *E. coli* NusG are intermolecular and transient. *Eingereicht*.

Die aufgeführten Experimente wurden von mir geplant und durchgeführt. Die Expression und Reinigung von *full-length* NusG und NusG-NTD wurden von Ulrich Scheckenhofer und von NusG-CTD von mir etabliert. Die NMR-Experimente wurden von mir mit Hilfe von Kristian Schweimer gemessen und ausgewertet. Das Manuskript wurde von Paul Rösch, Kristian Schweimer und mir verfasst.

6.6 Einzelarbeit F

Björn M. Burmann, Kristian Schweimer, Xiao Luo, Markus C. Wahl, Barbara L. Stitt, Max E. Gottesman und Paul Rösch (2010): A NusE:NusG Complex Links Transcription and Translation. *Science*, **328**, 501 – 504.

Die aufgeführten Experimente wurden von mir geplant und durchgeführt. Die NMR-Experimente wurden von mir mit Hilfe von Kristian Schweimer gemessen und ausgewertet. Interaktionstests zwischen NusG und NusE wurden von Xiao Luo durchgeführt. Rho für den Test der Interaktion mit NusG wurde von Barbara L. Stitt zur Verfügung gestellt. Das Manuskript wurde von Paul Rösch, Kristian Schweimer, Markus C. Wahl, Max E. Gottesman und mir verfasst.

Einzelarbeit A

Björn M. Burmann, Augusto Uc-Mass, Kristian Schweimer, Max E. Gottesman und Paul Rösch:

The Y39A Mutation of HK022 Nun Disrupts a *boxB* Interaction but Preserves Termination Activity. *Biochemistry* **47**, 7335 – 7341 (2008)

The Y39A Mutation of HK022 Nun Disrupts a *boxB* Interaction but Preserves Termination Activity[†]

Björn M. Burmann,[‡] Augusto Uc-Mass,[§] Kristian Schweimer,[‡] Max E. Gottesman,[§] and Paul Rösch^{*‡}

Department of Biopolymers and Research Center for Bio-Macromolecules, Universität Bayreuth, Universitätsstrasse 30, 95447 Bayreuth, Germany, and Department of Microbiology and Institute of Cancer Research, Columbia University Medical Center, New York, New York 10032

Received March 14, 2008; Revised Manuscript Received May 7, 2008

ABSTRACT: Coliphage HK022 Nun protein targets phage λ *nut boxB* RNA and acts as a transcriptional terminator, counteracting the phage λ N protein, a suppressor of transcription termination. Both Nun and N protein interact directly with RNA polymerase, and Nun competes with N protein for *boxB* binding and prevents superinfection of *Escherichia coli* HK022 lysogens by λ . Interaction of Trp18 of λ N and A7 of *boxB* RNA in the N–*boxB* complex is essential for efficient antitermination. We found that the corresponding Nun mutation, Nun Y39A, disrupts the interaction between the aromatic ring of Y39 and A7, but the mutant retains *in vivo* termination activity. Stabilization of the complex by interaction of A7 with an aromatic amino acid is thus less important for Nun activity than it is for N activity. Structural investigations show similar binding of mutant and wild-type (wt) Nun protein to *boxB* RNA. The dissociation constants of the wt Nun(20–44)–*boxB* and mutant Nun(20–44)–*boxB* complex as well as the structures of the *boxB* RNA in both complexes are identical.

Coliphage HK022 Nun protein blocks superinfection of HK022 lysogens by the related phage λ (*1, 2*). Both phage λ N protein and HK022 Nun protein interact with the *Escherichia coli* transcription elongation complex (TEC)¹ which includes RNA polymerase (RNAP) and *E. coli* proteins NusA, NusB, NusE (S10 ribosomal subunit), and NusG (*3, 4*). The consequence of this interaction is opposite for the two proteins: λ N suppresses transcription termination (*1, 5*), whereas HK022 Nun promotes it (*6*). Nun attaches to the *boxB* RNA sequence, a 15-mer hairpin structure with a purine-rich pentaloop (*7, 8*), of two *cis*-acting elements, λ *nutR* and λ *nutL*, of the λ nascent transcript. The sequence of *nutL boxB* differs from the sequence of *nutR boxB* by a single G-to-A substitution. The 107-amino acid protein Nun contains an amino-terminal domain with an as-yet-unknown function, followed by an arginine-rich motif (ARM) that binds *boxB* RNA (*9*), and a C-terminal putative DNA/RNAP interaction domain (*10*).

Both λ N and HK022 Nun bind to *boxB* RNA with similar affinities in the low nanomolar range via their ARM (λ N

ARM, QTRRRERRAEKQ; HK022 Nun ARM, RDRRRI-ARWEKR) (*11–13*). Similar to the N(1–36) peptide, the Nun(20–44) peptide containing the ARM folds into a bent α -helix upon complex formation, and *boxB* RNA attains a GNRA tetraloop formation with an extruded A7 (*12–14*). S24 and R28 intercalate with bases of the 5' stem, and it was proposed that the structure is further stabilized by π – π interaction between Y39 and *boxB* A7 which is observed in the λ N–*boxB* RNA complex W18–*boxB* A7 interaction (*12–14*). Our recent studies, however, show this amino acid–base stacking to be required neither for binding nor for correct RNA folding into the canonical GNRA tetraloop conformation that is attained by the pentaloop on λ N–HK022 Nun ARM binding. For wild-type (wt) λ N, Xia et al. (*15, 16*) observed that the stacked conformation is only transiently populated and an equilibrium between the stacked and open conformation on the picosecond time scale was observed. However, the requirement for an aromatic amino acid at this position for efficient antitermination *in vivo* and *in vitro* was determined (*15, 17, 18*). The Nun Y39A mutant shows termination activity *in vivo* despite the lack of the Nun Y39–*boxB* A7 interaction, in striking difference to λ N which requires the W18–*boxB* A7 base interaction for full activity. Combining fluorescence equilibrium titrations, NMR, and molecular dynamics (MD) simulations, we were able to relate structural features of the protein with its physiological termination activity *in vivo*.

MATERIALS AND METHODS

Sample Preparation. Unlabeled 15-nucleotide *nutR boxB* RNA was synthesized by *in vitro* transcription using T7 polymerase, a synthetic DNA template (5'-GCCCTTTTTCAGGGCTATAGTGAGTCGTATTA-3', Biomers, Ulm, Ger-

[†] This project was supported by a grant to P.R. from the Deutsche Forschungsgemeinschaft DFG (Ro617/16-1). M.E.G. is supported by NIH Grant GM37219.

* To whom correspondence should be addressed: Research Center for Bio-Macromolecules, Universität Bayreuth, Universitätsstr. 30, 95440 Bayreuth, Germany. Phone: +49 921 553540. Fax: +49 921 16490459. E-mail: roesch@unibt.de.

[‡] Universität Bayreuth.

[§] Columbia University Medical Center.

¹ Abbreviations: ARM, arginine-rich motif; *nut*, N utilization site; Nus, N utilization substance; MD, molecular dynamics; COSY, correlated spectroscopy; EOP, efficiency of plating; NOESY, nuclear Overhauser enhancement spectroscopy; RNAP, RNA polymerase; rmsd, root-mean-square deviation; TEC, transcription elongation complex; TOCSY, total coherence spectroscopy.

many), and unlabeled nucleotide triphosphates. RNA was purified as described previously (13). Freeze-dried RNA was resuspended in NMR buffer [50 mM potassium phosphate and 40 mM NaCl (pH 6.4)], heated for 5 min at 95 °C, cooled for refolding, dialyzed against water, and freeze-dried for storage. This RNA was directly dissolved in NMR buffer for experiments. 3'-(6-Fam)-labeled *nutR boxB* for fluorescence measurements was obtained from Biomers and used according to the manufacturer's instructions.

Nun(20–44) (RGLTSRDRRRRIARWEKRIAYALKNG) and Nun(20–44) Y39A were purchased from PANATecs (Tübingen, Germany), dialyzed against water for desalting, and freeze-dried. Nun(20–44) Y39A–*nutR boxB* RNA samples were generated by adding small volumes of a concentrated RNA solution to the peptide. For further concentration, the complex was freeze-dried and resuspended in a H₂O/D₂O mixture. Part of the peptide aggregated upon binding to the RNA as observed for wt Nun (12), and concentrations were thus limited to less than 500 μM.

Fluorescence Measurements. Extrinsic fluorescence measurements with 3'-6-carboxyfluorescein (6-Fam)-labeled *nutR boxB* RNA were performed in NMR buffer in a volume of 1 mL using a 10 mm × 4 mm quartz cuvette (Hellma, Müllheim, Germany) with an L-format Jobin-Yvon (Edison, NJ) Horiba Fluoromax fluorimeter. The excitation wavelength was 492 nm, and the emission intensity was detected at 516 nm applying a 500 nm cutoff filter. The slit widths were 9 and 7 nm for excitation and emission, respectively, and all titration measurements were taken at 25 °C with the fluorescence-labeled RNA at 50 pM. Stock peptide solutions contained 1 μM Nun and Nun(20–44) Y39A. Following sample equilibration, 100 data points with an integration time of 0.1 s were collected for each titration step.

Data Fitting. Data were fitted to a two-state binding equation to determine the equilibrium dissociation constant (K_d) using standard software:

$$F = F_{\min} - (F_{\max} - F_{\min}) \left[\frac{\sqrt{(K_d + [P]_0 + [RNA]_0)^2 - 4[P]_0[RNA]_0}}{2[RNA]_0} \right] \quad (1)$$

where F is the fluorescence intensity, F_{\max} and F_{\min} are the signal intensities of the bound and unbound form, respectively, and $[P]_0$ and $[RNA]_0$ are the total protein and RNA concentration, respectively.

For displacement experiments, labeled RNA and Nun peptide at 10 nM were used, and the absolute concentrations were significantly higher than the determined dissociation constant. Therefore, the titration starts at a 1:1 ratio of Nun to labeled RNA with negligible concentrations of free protein and RNA in solution, and the displacement titration was evaluated with (19)

$$F = F_{\max} - (F_{\max} - F_{\min}) \left[\frac{\sqrt{(K_r([P]_0 + [RNA]_0))^2 - 4(K_r - 1)K_r[P]_0[RNA]_0}}{[P]_0(K_r - 1)} \right] \quad (2)$$

where F_{\max} is the relative fluorescence intensity at the beginning of the titration, F_{\min} is the intensity under saturating conditions of the unlabeled RNA, and K_r is the relative affinity of the unlabeled and labeled RNA ($K_r = K_1/K_2$, where K_1 is the dissociation constant of labeled RNA and K_2 is the dissociation constant of unlabeled RNA).

NMR Measurements. All NMR experiments were recorded at either 298 K (one-dimensional experiments) or 303 K (two-dimensional experiments) on Bruker DRX 600 MHz and AV 700 MHz spectrometers with triple-resonance probes equipped with pulsed field gradient capabilities. For resonance assignments, correlated spectroscopy (COSY), total coherence spectroscopy (TOCSY), and nuclear Overhauser enhancement spectroscopy (NOESY) experiments were performed using standard techniques for recording and water suppression (20). All NMR data were analyzed with NMRView version 5.2.2 (21) and in house routines.

All our experimental data clearly indicated that the fold of the *boxB* RNA in the Nun Y39A complex is identical to the fold of the wt Nun–*boxB* [Protein Data Bank (PDB) entry 1HJI] complex and thus virtually identical to the fold of *boxB* in the λ N(1–36)–*boxB* (PDB entry 1QFQ) complex (12, 13), rendering it plausible to use this *boxB* RNA structure as a fixed template for all MD calculations.

All structure calculations were performed using a modified ab initio simulated annealing protocol with an extended version of Xplor-NIH 1.2.1 (22, 23). The calculation strategy as described previously (24) included floating assignment of prochiral groups (25), a conformational database potential term (26), and a reduced presentation for nonbonded interactions for part of the calculation (24). The protocol was identical to that used previously (11). For analysis, PROCHECK-NMR (27) was used, and for graphical presentations, PyMol (28) was used.

In Vivo Assays. Plasmid *pTrc99* (Amersham Bioscience) is a ColE1 plasmid encoding ampicillin resistance. Plasmid *pTrc-Nun* is *pTrc99* carrying *nun* under *tac* promoter control (Gottesman laboratory collection), resulting in a general test system for Nun activity. Plasmid *pTrc-Nun Y39A* was constructed by introducing a mutation into codon 39 of the *nun* gene cloned in *pTrc-Nun* using the Quickchange site-directed mutagenesis kit (Stratagene). The mutation substitutes a tyrosine residue at position 39 for alanine. The following primers were used: Y39Af, TGG GAA AAA AGG ATA GCA GCC GCA TTA AAA AAC GGT GTG; and Y39Ar, CAC ACC GTT TTT TAA TGC GGC TGC TAT CCT TTT TTC CCA.

Strains. Strains used in this study were *E. coli* W3102 (N99, NIH collection) and its derivatives *N99lacZXA21* (Gottesman laboratory collection), *N99 λcI857-pL-nutL-N:lacZ* (29), and *N99lacZXA21 λcI857-pR-cro-nutR-cII:lacZ* (from N. Costantino and D. Court).

Efficiency of λ Plating (EOP). Fresh overnight cultures of N99 transformed with pTrc99, pTrc-Nun, and pTrc-Nun Y39A were poured atop agar on LB plates. EOPs were determined by spotting dilutions of λ and incubation overnight at 37 °C.

Termination Efficiency. Log phase cultures of strains N99 *lacZXA21*, *N99 λcI857-pL-nutL-N:lacZ*, and *N99 lacZXA21 λcI857-pR-cro-nutR-cII:lacZ* transformed with pTrc99, pTrc-Nun, and pTrc-Nun Y39A were heated from 32 to 42 °C and incubated for 1 h. Samples were assayed for β-galactosidase activity as described previously (30). The shift to 42 °C inactivates the λcI857 repressor, permitting transcription from the λpL and λpR promoters. The percent of read-through was calculated as described by Kim et al. (31).

MD Simulations. The Amber 9 program package (32) and the *ff03* force field (33, 34) were used for the simulations of

Table 1: Nun Y39A Is as Efficient as wt Nun for λ Exclusion and Transcription Termination^a

plasmid	read-through (%)		
	EOP ^b	<i>pL-nutL-N:lacZ</i> ^c	<i>pR-cro-nutR-tR1-cII:lacZ</i> ^c
pTrc99	1	100	100
pTrc-Nun	10 ⁻⁶	1	3
pTrc-Nun Y39A	10 ⁻⁶	9	7

^a Induction is for 1 h at 42 °C. λ -Galactosidase assays are as described previously (29). All data are averages of two independent experiments. ^b Strain W3102 transformed with the indicated plasmids was spotted with λ at 37 °C. ^c Nun sensitive fusions are controlled by the temperature sensitive cI857 repressor.

the wt Nun(20–44)–*nutR boxB* and Nun(20–44) Y39A–*nutR boxB* complexes. The simulation for the wt complex was based on the NMR structure of the complex (PDB entry 1HJI). For the *in silico* mutation of the peptide, the Loopy program from the Jackal suite (http://wiki.c2b2.columbia.edu/honiglab_public/index.php/Software:Jackal) was used to alter the wt PDB entry and to provide a correct insertion of A39.

The RNA–peptide complexes were solvated in a TIP3P water box (35) with dimensions of 60 Å × 60 Å × 60 Å, and sodium counterions were added *in silico* for neutralization of the system. Calculations were performed at 298 K and an external pressure of 1 atm. Under these conditions, the systems were minimized and equilibrated using SANDER. Initially, the whole system was minimized for 1000 steps, and the water molecules and counterions were relaxed around the fixed solute with a 100 ps MD run. MD production runs 20 ns in duration were then performed for both systems. The MD data were analyzed by using PTRAJ. Root-mean-square deviation (rmsd) calculations of the atomic coordinates were referenced to the first calculated structure.

RESULTS AND DISCUSSION

Nun Y39A Is Active in Vivo. Phage HK022 Nun protein excludes superinfection of λ by terminating the λ pL and λ pR early transcripts just distal from the *nut* sites. We tested the *in vivo* activity of Nun Y39A by determining the efficiency of plating (EOP) of λ on a lawn of *E. coli* carrying plasmid *pTrc-Nun Y39A*. We found that λ forms plaques on cells expressing wild-type Nun or Nun Y39A with very low efficiency (<10⁻⁶) compared to the control strain carrying empty vector (in Table 1, compare rows 2 and 3 with row 1). These results show that Nun Y39A is as efficient as wt Nun at excluding λ . We then tested Nun Y39A for termination in two λ *nut-lacZ* fusions (see Materials and Methods). The level of β -galactosidase activity provides a quantitative measure of Nun termination efficiency at *nutL* and *nutR*. Our data show that the percent read-through with Nun Y39A is only slightly higher than with wt Nun protein (in Table 1, compare columns 2 and 3). In conclusion, our data show no significant differences between the activities of Nun Y39A and wt Nun, suggesting that the binding of Nun Y39A to *boxB* should be equivalent to that of wt Nun as far as termination activity is concerned.

Mutation of equivalent amino acids affected HK022 Nun and λ N activity to a different degree (Table 2) (17, 18). Amino acids whose variation in λ N led to a loss of RNA binding activity according to gel mobility shift assays caused only a minor loss of termination activity of HK022 Nun, with the most notable effect caused by Nun mutations R25A,

Table 2: Effect of Nun Mutations on Antitermination Activity

	<i>pL-cro-nutR-lacZ</i> ^a	λ exclusion ^b	λ <i>nin</i> exclusion ^{b,c}
Nun ⁺	97	+	+
vector	0	–	–
R25A	71	+	+
R27A	84	+	+
A31L	93	+	+
R32A	83	+	+
K35A	93	+	+
R36A	91	+	+
Y39A	93	+	+
N43A	96	+	+

^a Values represent the percent termination based on β -galactosidase assays. ^b Exclusion was determined as described in Materials and Methods. ^c λ *nin* is deleted for terminators between λ genes P and Q and does not require N for growth. Thus, the Nun mutants are not merely competing with λ N for *boxB* RNA binding but are actively terminating transcription.

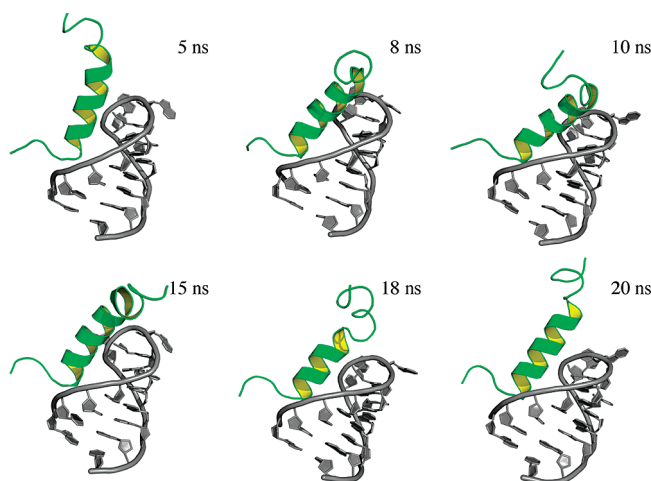


FIGURE 1: Snapshots taken from the MD simulation of the Nun Y39A–RNA complex showing a high degree of flexibility for amino acids 31–44 of the Nun Y39A peptide (green). Typical features of the RNA (gray) are the helical turn in the stem region and the GNRA tetraloop with extruded base A9. The RNA is slightly flexible in the loop region.

R27A, and R32A (71, 84, and 83% termination efficiency, respectively) in the contact region of Nun and the RNA stem (12). λ N activity is thus more critically dependent on key amino acid variations than Nun activity.

The Carboxy Terminus of Nun(20–44) Y39A Is Highly Flexible in MD Simulations. MD simulations of the wt Nun(20–44)–*boxB* RNA and Nun(20–44) Y39A–*boxB* RNA complexes showed that the conformation of Nun(20–44) Y39A in the complex is defined well only from S24 to I30, whereas the wt complex is stable and defined well over the whole simulation time of 20 ns. The carboxy-terminal helix of the peptide, R32–G44, possessed a high degree of flexibility without preferential orientation but retained helical structure (Figure 1). Other indicators of flexibility of the peptide in the complex are the rmsd values of both simulations (Figure 2A), indicating the mutant peptide in the complex to be more flexible than the wt peptide. The variation of the distance between Y39 and RNA A7 for the wt complex compared to the mutant A39 and RNA A7 distance rules out hydrophobic interaction between the three alanines (38–40) and RNA A7 as the distance between H β atoms of these alanines and RNA A7 H8 is greater than 10 Å.

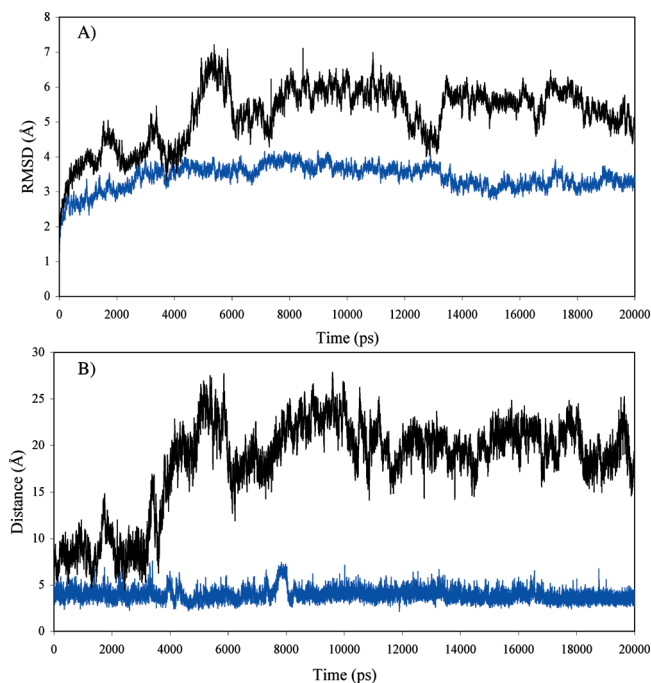


FIGURE 2: (A) rmsd values for Nun (blue) and Nun Y39A (black) as a function of simulation time referenced to the first calculated structure. (B) Distance between Tyr39 H δ 1 and A7 H8 for Nun (blue) and Ala39 H β 2 and A7 H8 for NunY39A (black) as a function of simulation time.

Surprisingly, the structure of the RNA remains stable in both simulations, retaining the helical turn in the stem region as well as the GNRA tetraloop with A9 extruded. Thus, the equilibrium between folded and unfolded RNA is shifted toward the folded species by binding of either peptide, although the RNA loop region is more flexible in the complex with the mutant peptide.

NunY39A Forms a Well-Defined Complex with the *boxB* RNA. One-dimensional NMR spectra show directly the binding of the peptide to the RNA in the imino proton region. For free RNA, only the imino proton resonances of G12, G13, and G14 were observed. Upon addition of Nun(20–44) Y39A, these resonances shifted, and the imino proton resonances of U5 and G6 as well as the indole NH resonance of W33 could be detected (compare Figure 3B to wt in Figure 3A). This increase in number and the change in the position of these resonances are typical signs of the stabilization of *boxB* by these peptides. Both Nun complexes resulted in the same number and pattern of imino proton signals between 10 and 14 ppm, indicating that the mutant peptide induces RNA stem structure similar to the structure induced by the wt peptide, and appearance of the G6 resonance suggests the formation of the sheared G6–A10 base pair that leads to the formation of the GNRA tetraloop. The emergence of the U5 imino resonance in either spectrum relates to the stabilization of the apical U5–A11 base pair. These results indicate that the binding of the mutant peptide to the *boxB* RNA follows the same general pattern in the stem region as the wt peptide.

The wt Nun–*boxB* Complex Is Marginally More Stable Than the Nun Y39A–*boxB* Complex. To determine the contribution of the tyrosine base stacking interaction to complex stability, the K_d values were determined for the mutant and wt peptide with fluorescence titrations at 516

nm employing 3'-(6-Fam)-labeled *nutR boxB* RNA. K_d values were 2.4 ± 0.1 and 6.1 ± 0.3 nM for the peptide–RNA complex at 25 °C for the wt and mutant peptide, respectively, suggesting the π – π interaction between the peptide and RNA contributes only marginally to complex stability (Figure 4). No results were obtained with 5'-(6-Fam)-labeled *boxB* RNA, indicating that the peptides did not directly interact with the fluorescence label. Additionally, a displacement titration resulted in a K_d of 3.8 ± 1.0 nM for unlabeled *boxB* in the wt complex, in the range of the K_d for Nun and the 3'-(6-Fam)-labeled *nutR boxB* RNA (Figure 5).

Structure of the Variant HK022 Nun–*boxB* Complex. Standard homonuclear two-dimensional NMR experiments (COSY, TOCSY, and NOESY) proved to be sufficient for obtaining sequence-specific resonance assignments for Nun(20–44) Y39A, and several NOESY cross-peaks in the backbone amide–amide region indicating a helical fold for large parts of the peptide could be assigned (Table 3). In particular, for amino acids T23–R29 and I37–L42, helix-typical i – $i + 3$ and i – $i + 4$ NOEs could be observed, but only very few nonsequential NOEs could be detected from amino acid I30 to R36, indicating structural flexibility in this part of the peptide. For the C- and N-terminal amino acids, no nonsequential NOEs could be detected. The ribose resonances for RNA A9 showed the exceptional downfield shift that was already observed for the wt complexes of HK022 Nun and λ N. Consistent with this feature, the H1' $_i$ –H6/H8 $_{i+1}$ NOEs that are typically used for the assignment of the ribose spin system and that of the following base (36) are missing between A8 and A9 as well as between A9 and A10. Overall, the chemical shifts of the RNA protons of the mutant complex are very similar to the shifts observed for the wt complexes. For further structure calculations, we used the *nutR boxB* RNA conformation from the N36–*boxB* RNA complex as calculated in ref 13 as a starting point, in analogy to ref 12.

Thirteen intermolecular NOEs could be identified unambiguously in the NOESY spectra. For the β -protons of S24, NOEs with RNA C2 and C3 were observed. Additionally, NOEs between the side chain protons of R28, C4 H5 and U5 H5, could be assigned, suggesting that the interaction between Nun(20–44) Y39A and *boxB* RNA is restricted to S24, R28, and the stem region of the RNA (Figure 6), in good agreement with the observations for wt Nun(20–44) (12) and mutational studies of the λ N peptide (15, 18).

A PROCHECK-NMR (27) analysis of the Nun Y39A–*boxB* RNA complex shows that 80.0% of the residues of the 20 accepted structures are found in the most favored regions and an additional 19.6% in the allowed regions of the Ramachandran plot (Table 3). For Nun(20–44) Y39A, formation of an α -helix for residues S24–I30 as well as an additional α -helical turn for residues A38–L41 was found, in good agreement with the wt complex structure, in which Nun(20–44) forms a bent α -helix for residues S24–N43, with the kink at A31 and R32. Clearly, the α -helix for Nun(20–44) Y39A from residue S24 to I30 is stabilized by its interaction with *boxB* RNA, and the loss of secondary structure in the carboxyl-terminal part is due to the lack of stabilization by the π – π interaction between Y39 and RNA A7 that is present in the wt complex. The α -helical turn for residues A38–L41 is independently induced by the helix favoring three alanines and one leucine in this sequence. The

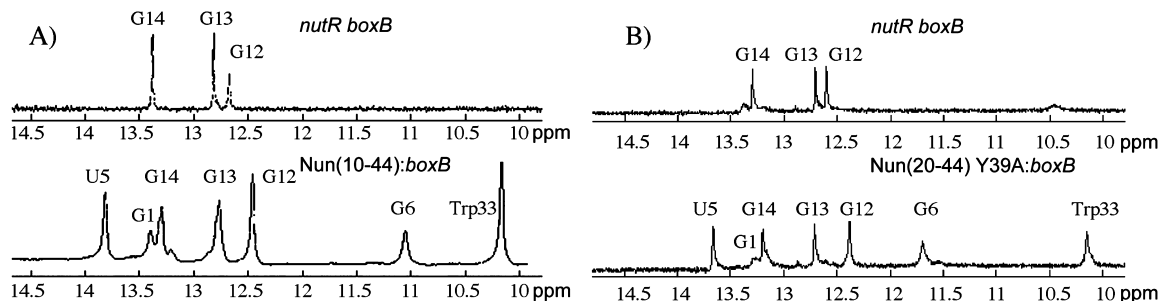


FIGURE 3: (A) One-dimensional NMR spectra of the imino proton region of *nutR boxB* RNA and the Nun(10–44)–*boxB* complex (11). (B) One-dimensional NMR spectra of the imino proton region of *nutR boxB* RNA and the Nun(20–44) Y39A–*boxB* complex. All spectra were recorded in a H₂O/D₂O mixture (9:1), 40 mM NaCl, and 50 mM potassium phosphate (pH 6.4) at 298 K.

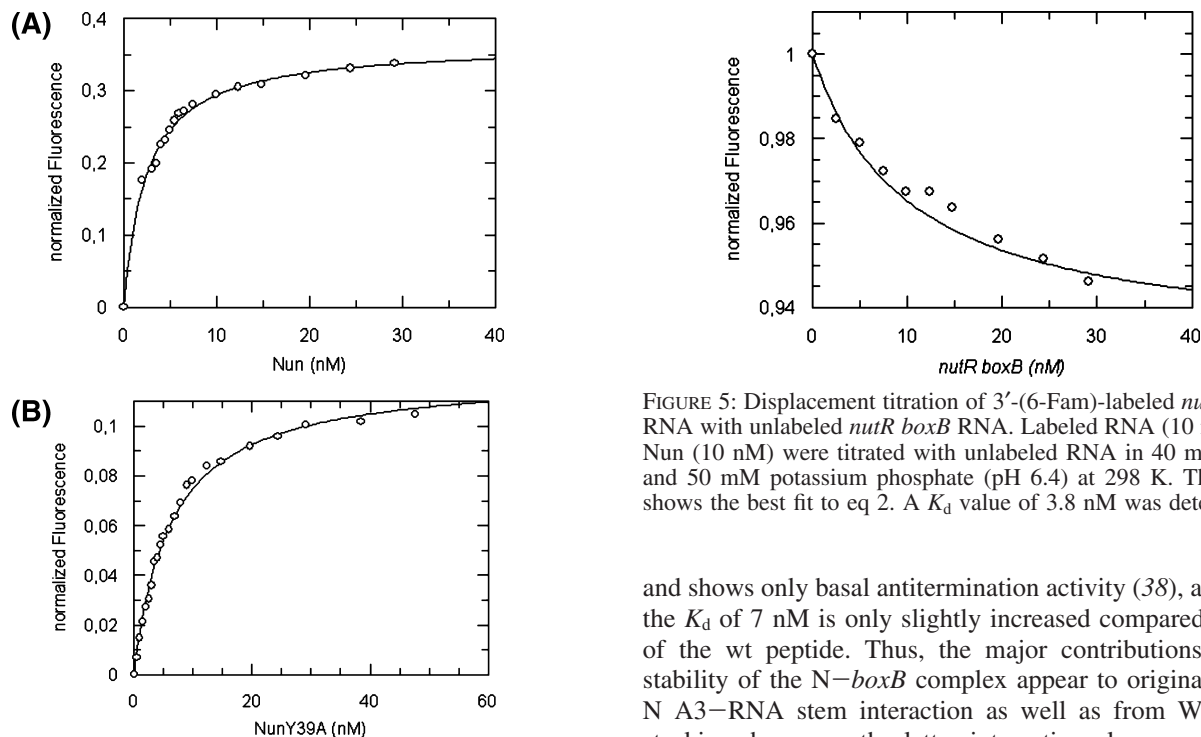


FIGURE 4: Fluorescence equilibrium titrations with (A) Nun and (B) Nun Y39A. 3'-(6-Fam)-labeled *nutR boxB* RNA (50 pM) was titrated with each peptide in 40 mM NaCl and 50 mM potassium phosphate (pH 6.4) at 298 K. The curves show the best fit to eq 1. A K_d value of 2.4 nM was determined for Nun and a K_d value of 6.1 nM for Nun Y39A.

experimental results thus clearly confirm the MD simulations, that is, formation of an α -helix for residues S24–I30 and higher flexibility in the carboxyl-terminal part with a short α -helical turn for residues A38–L41.

Comparison of λ N-*boxB* and HK022 Nun-*boxB* Complexes. The conformations of *boxB* RNA complexes of λ N and HK022 Nun ARM peptides are highly similar, and the K_d values for both are in the low nanomolar range as confirmed here [2.4 ± 0.1 and 1.3 ± 0.4 nM (11) for Nun and N, respectively]. This difference can be pinpointed to the presence of S23 in Nun at the position equivalent to A3 in λ N; λ N A3S showed a decrease to 80% compared to the wt affinity in gel mobility shift assays (18). Thus, A3 is key for the stability of the N ARM peptide-*boxB* complex.

The increase in the K_d for the Nun Y39A-*boxB* complex is negligible, whereas for the N W18A mutant, no RNA binding could be detected in gel shift experiments (18). The λ N double mutant E14R15 lacks the W18–A7 interaction

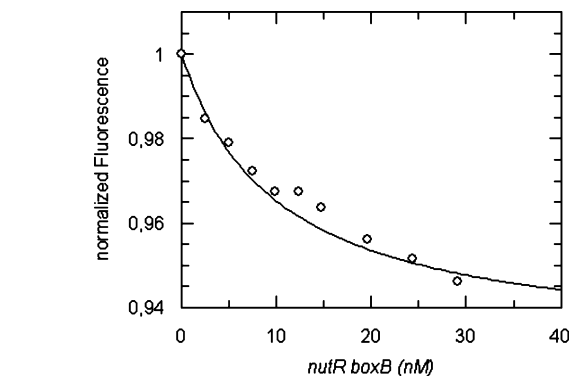


FIGURE 5: Displacement titration of 3'-(6-Fam)-labeled *nutR boxB* RNA with unlabeled *nutR boxB* RNA. Labeled RNA (10 nM) and Nun (10 nM) were titrated with unlabeled RNA in 40 mM NaCl and 50 mM potassium phosphate (pH 6.4) at 298 K. The curve shows the best fit to eq 2. A K_d value of 3.8 nM was determined.

and shows only basal antitermination activity (38), although the K_d of 7 nM is only slightly increased compared to that of the wt peptide. Thus, the major contributions to the stability of the N-*boxB* complex appear to originate from N A3–RNA stem interaction as well as from W18–A7 stacking; however, the latter interaction also seems to be crucial for functionality. In contrast, the HK022 Nun-*boxB* complex is neither functionally nor energetically dependent on the corresponding stacking interaction. This follows the pattern in which tyrosine stacking was reported to be energetically less favorable than tryptophan stacking (39). Thus, weakened stacking interaction and the serine-for-alanine substitution in Nun could explain the slightly higher K_d of the Nun complex, although the extent of the K_d increase seems far too small. Thus, the tyrosine-for-tryptophan and serine-for-alanine substitution must to some extent be offset by other amino acid–base interactions. The availability of an EOP test for Nun mutants and the combination of the test results with MD calculations are expected to shed further light on the details of this peptide–RNA interaction.

Even mutations of amino acids that are involved in binding to the RNA stem region do not lead to a complete loss of Nun termination activity, which, on a functional level, could be explained by the fact that for efficient termination, the Nun-*boxB* complex needs to persist only for the time span needed by the TEC to read through ~ 100 bp; the λ N-*boxB* complex, however, has to persist for the time span the TEC needs to read through several kilo base pairs. This significant difference could explain the importance of the W18–A7 π - π interaction with respect to functionality as it may help

Table 3: Structural Statistics

Experimentally Derived Restraints	
no. of distance restraints	
NOEs	202
intraresidual	0
sequential	98
medium-range	91
long-range	0
intermolecular	13
Restraint Violations	
average distance restraint violation	0.0075 ± 0.0024 Å
maximum distance restraint violation	<0.1 Å
Deviation from Ideal Geometry	
bond lengths	0.00225 ± 0.00003 Å
bond angles	0.489 ± 0.003°
Coordinate Precision ^a	
Arg20–Gly44	Ser24–Lys42
backbone heavy atoms, 2.12 Å	backbone heavy atoms, 1.53 Å
all heavy atoms, 3.10 Å	all heavy atoms, 2.57 Å
Arg20–Gly44 + <i>boxB</i> RNA	Ser24–Lys42 + <i>boxB</i> RNA
backbone heavy atoms, 2.56 Å	backbone heavy atoms, 2.10 Å
all heavy atoms, 3.55 Å	all heavy atoms, 3.07 Å
Ramachandran Plot Statistics ^b	
	80.0%, 18.2%, 1.4%, 0.5%

^a The precision of the coordinates is defined as the average atomic root-mean-square deviation between the accepted simulated annealing structures and the corresponding mean structure calculated for the given sequence regions. ^b Ramachandran plot statistics are determined by PROCHECK-NMR and given in the following order: most favored, additionally allowed, generously allowed, disallowed.

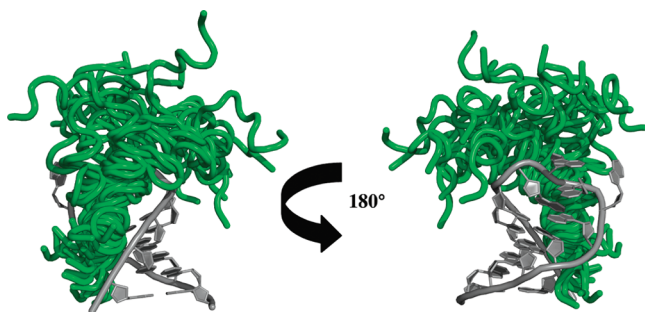


FIGURE 6: Overlay of the 20 lowest-energy structures of the Nun Y39A–*boxB* RNA complex. The peptide structure (green) is well-defined at the RNA (gray) interface but is highly flexible after the bend between amino acids 31 and 33. Both regions of the peptide are helical. Structural calculations were performed with a total number of 189 intramolecular and 13 intermolecular distance restraints. All structures exhibited NOE violations of <0.1 Å.

to add that small amount of binding energy necessary to keep the λ N–*boxB* complex stable long enough for antitermination, but which is not necessary for the short lifetime required for the HK022 Nun–*boxB* complex to exert its activity.

ACKNOWLEDGMENT

We thank Ulrike Persau and Britta Zimmermann for RNA preparations. We thank J. Pan for construction of *pTrc-Nun Y39A* and N. Costantino and D. Court for the *nut-lacZ* fusions.

REFERENCES

1. Robert, J., Sloan, S. B., Weisberg, R. A., Gottesman, M. E., Robledo, R., and Harbrecht, D. (1987) The remarkable specificity

- of a new transcription termination factor suggests that the mechanisms of termination and antitermination are similar. *Cell* 51, 483–492.
2. Robledo, R., Gottesman, M. E., and Weisberg, R. A. (1990) Lambda *nutR* mutations convert HK022 Nun protein from a transcription termination factor to a suppressor of termination. *J. Mol. Biol.* 212, 635–643.
3. Das, A. (1993) Control of transcription termination by RNA-binding proteins. *Annu. Rev. Biochem.* 62, 893–930.
4. Greenblatt, J., Nodwell, J. R., and Mason, S. W. (1993) Transcriptional antitermination. *Nature* 364, 401–406.
5. Chattopadhyay, S., Garcia-Mena, J., DeVito, J., Wolska, K., and Das, A. (1995) Bipartite function of a small RNA hairpin in transcription antitermination in bacteriophage lambda. *Proc. Natl. Acad. Sci. U.S.A.* 92, 4061–4065.
6. Oberto, J., Weisberg, R. A., and Gottesman, M. E. (1989) Structure and function of the *nut* gene and the immunity region of the lambdaoid phage HK022. *J. Mol. Biol.* 207, 675–693.
7. Franklin, N. C. (1985) Conservation of genome form but not sequence in the transcription antitermination determinants of bacteriophages λ , ϕ 21 and P22. *J. Mol. Biol.* 181, 75–84.
8. Lazinski, D., Grzadziska, E., and Das, A. (1989) Sequence-specific recognition of RNA hairpins by bacteriophage antiterminators requires a conserved arginine-rich motif. *Cell* 59, 207–218.
9. 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.
10. Watnick, R. S., Herring, S. C., Palmer, A. G., III, and Gottesman, M. E. (2000) The carboxyl terminus of phage HK022 Nun includes a novel zinc-binding motif and a tryptophan required for transcription termination. *Genes Dev.* 14, 731–739.
11. Van Gilst, M. R., Rees, W. A., Das, A., and von Hippel, P. H. (1997) Complexes of N antitermination protein of phage λ with specific and nonspecific RNA target sites on the nascent transcript. *Biochemistry* 36, 1514–1524.
12. Faber, C., Schärpf, M., Becker, T., Sticht, H., and Rösch, P. (2001) The structure of the coliphage HK022 Nun protein– λ –phage *boxB* RNA complex. Implications for the mechanism of transcription termination. *J. Biol. Chem.* 276, 32064–32070.
13. Schärpf, M., Sticht, H., Schweimer, K., Boehm, M., Hoffmann, S., and Rösch, P. (2000) Antitermination in bacteriophage λ . The structure of the N36 peptide–*boxB* RNA complex. *Eur. J. Biochem.* 267, 2397–2408.
14. Legault, P., Li, J., Mogridge, J., Kay, L. E., and Greenblatt, J. (1998) NMR structure of the bacteriophage λ N peptide/*boxB* RNA complex: Recognition of a GNRA fold by an arginine-rich motif. *Cell* 93, 289–299.
15. 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.
16. Xia, T., Wan, C., Roberts, R. W., and Zewail, A. H. (2005) RNA-protein recognition: Single-residue ultrafast dynamical control of structural specificity and function. *Proc. Natl. Acad. Sci. U.S.A.* 102, 13013–13018.
17. Franklin, N. C. (1993) Clustered arginine residues of bacteriophage λ N protein are essential to antitermination of transcription, but their locale cannot compensate for *boxB* loop defects. *J. Mol. Biol.* 231, 343–360.
18. Su, L., Radek, J. T., Hallenga, K., Hermanto, P., Chan, G., Labeots, L. A., and Weiss, M. A. (1997) RNA recognition by a bent α helix regulates transcriptional antitermination in phage λ . *Biochemistry* 36, 12722–12732.
19. Müller, B., Restle, T., Reinstein, J., and Goody, R. S. (1991) Interaction of fluorescently labeled dideoxynucleotides with HIV-1 reverse transcriptase. *Biochemistry* 30, 3709–3715.
20. Cavanagh, J., Fairbrother, W. J., Palmer, A. G., III, Rance, M., and Skelton, N. J. (2006) *Protein NMR spectroscopy: principles and practice*, Academic Press, San Diego.
21. 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.
22. Schwieters, C. D., Kuszewski, J. J., and Clore, G. M. (2006) Using Xplor-NIH for NMR molecular structure determination. *Prog. Nucl. Magn. Reson. Spectrosc.* 48, 47–62.
23. 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.

24. Nilges, M., Clore, G. M., and Gronenborn, A. M. (1988) Determination of three-dimensional structures of proteins from interproton distance data by dynamical simulated annealing from a random array of atoms. *FEBS Lett.* **239**, 129–136.
25. Holak, T. A., Nilges, M., and Oschkinat, H. (1989) Improved strategies for the determination of protein structures from NMR data: The solution structure of acyl carrier protein. *FEBS Lett.* **242**, 218–224.
26. Kuszewski, J., Gronenborn, A. M., and Clore, G. M. (1996) Improving the quality of NMR and crystallographic protein structures by means of a conformational database potential derived from structure databases. *Protein Sci.* **5**, 1067–1080.
27. Laskowski, R. A., Rullmann, J. A., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) AQUA and PROCHECK-NMR: Programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR* **8**, 477–486.
28. DeLano, W. L. (2002) *PyMOL*, DeLano Scientific, Palo Alto, CA.
29. Wilson, H. R., Kameyama, L., Zhou, J. G., Guarneros, G., and Court, D. L. (1997) Translational repression by a transcriptional elongation factor. *Genes Dev.* **11**, 2204–2213.
30. Miller, J. H. (1992) *A Short Course in Bacterial Genetics: A Laboratory Manual for Escherichia coli and Related Bacteria*, Cold Spring Harbor Laboratory Press, Plainview, NY.
31. Kim, H. C., Washburn, R. S., and Gottesman, M. E. (2006) Role of *E. coli* NusA in phage HK022 Nun-mediated transcription termination. *J. Mol. Biol.* **359**, 10–21.
32. Case, D. A., Darden, T. A., Cheatham, T. E., Simmerling, C. L., Wang, J., Duke, R. E., Luo, R., Merz, K. M., Pearlman, D. A., Crowley, M., Walker, R. C., Zhang, W., Wang, B., Hayik, S., Roitberg, A., Seabra, G., Wong, K. F., Paesani, F., Wu, X., Brozell, S., Tsui, V., Gohlke, H., Yang, L., Tan, C., Mongan, J., Hornak, V., Cui, G., Beroza, P., Mathews, D. H., Schafmeister, C., Ross, W. S., and Kollman, P. A. (2006) *Amber 9*, University of California, San Francisco.
33. Duan, Y., Wu, C., Chowdhury, S., Lee, M. C., Xiong, G., Zhang, W., Yang, R., Cieplak, P., Luo, R., Lee, T., Caldwell, J., Wang, J., and Kollman, P. (2003) A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *J. Comput. Chem.* **24**, 1999–2012.
34. Lee, M. C., and Duan, Y. (2004) Distinguish protein decoys by using a scoring function based on a new AMBER force field, short molecular dynamics simulations, and the generalized born solvent model. *Proteins* **55**, 620–634.
35. Price, D. J., and Brooks, C. L., III (2004) A modified TIP3P water potential for simulation with Ewald summation. *J. Chem. Phys.* **121**, 10096–10103.
36. Johnson, N. P., Baase, W. A., and von Hippel, P. H. (2005) Low energy CD of RNA hairpin unveils a loop conformation required for λ N antitermination activity. *J. Biol. Chem.* **280**, 32177–32183.
37. Varani, G., and Tinoco, I., Jr. (1991) RNA structure and NMR spectroscopy. *Q. Rev. Biophys.* **24**, 479–532.
38. Xia, T., Becker, H. C., Wan, C., Frankel, A., Roberts, R. W., and Zewail, A. H. (2003) 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.
39. Kobayashi, H., Kato, J., Morioka, H., Stewart, J. D., and Ohtsuka, E. (1999) Tryptophan H33 plays an important role in pyrimidine (6–4) pyrimidone photoproduct binding by a high-affinity antibody. *Protein Eng.* **12**, 879–884.

BI8004347

Einzelarbeit B

Irena Matečko, Björn M. Burmann, Kristian Schweimer, Hubert Kalbacher, Jürgen Einsiedel, Peter Gmeiner und Paul Rösch:

Structural Characterisation of the *E. coli* Heat Stable Enterotoxin STh. *The Open Spectroscopy Journal* **2**, 34 – 39 (2008)

Structural Characterisation of the *E. coli* Heat Stable Enterotoxin STh

Irena Matecko¹, Bjoern M. Burmann¹, Kristian Schweimer¹, Hubert Kalbacher², Jürgen Einsiedel³, Peter Gmeiner³ and Paul Rösch^{*,1}

¹University of Bayreuth, Department of Biopolymers & Research Center for Bio-Macromolecules, D-95440 Bayreuth, Germany

²PANATecs GmbH, Vor dem Kreuzberg 17, D-72070 Tübingen, Germany

³Department of Medicinal Chemistry, Friedrich-Alexander-University Erlangen - Nürnberg, Schuhstrasse 19, D-91052 Erlangen, Germany

Abstract: *E. coli* heat stable enterotoxin STa is an agonist of the membrane guanylate cyclase C whose endogenous ligands are the peptide hormones guanylin and uroguanylin. Whereas these peptides contain only two disulfide bonds, STa is stabilized by one additional disulfide bridge. We chemically synthesized the enterotoxin STh that originates from the *E. coli* strain found in humans, and we determined its structure and its dynamics by nuclear magnetic resonance spectroscopy and molecular dynamics calculations. Chemical synthesis clearly proved successful and resulted in the formation of the native disulfide bonds. The endogenous ligands guanylin and uroguanylin show the same general structural features and dynamics properties as the enterotoxin.

Keywords: Enterotoxin, guanylyl cyclase, STh, STa, guanylin, uroguanylin.

INTRODUCTION

Many bacterial pathogens synthesize toxins that serve as virulence factors. Recently, these toxins became a topic of interest as a medication [1-3], inactive toxin components (toxoids) were suggested to be used as a vaccine [4], toxins were used as tools to elucidate the complex events during signal transduction [5], even as tumor markers and potential therapeutics in the treatment of colorectal and breast cancer [6-8].

Enterotoxigenic *E. coli* bacteria (ETEC) produce two forms of heat-stable enterotoxins: STa (or STI) and STb (or STII) [9,10]. These toxins cause acute and secretory diarrhea in humans, known as traveler's disease. In developing countries, this type of diarrhea is a major cause of death of infants [11]. STa consists of two subtypes that differ slightly in amino acid sequence and that are, for historic reasons, called STh (originally thought to occur in human *E. coli* strains only) and STp (originally thought to occur in porcine *E. coli* strains only).

STh is expressed as a precursor protein of 72 amino acids and it is cleaved twice before it is secreted as the mature 19 amino acid toxin [12-15]. The toxic domain of STh is located in its carboxy-terminal region, between C6-C18, and it is highly conserved within the whole toxin family [5]. The 6 cysteins in this domain are arranged in three disulfide bridges, C6-C10, C7-C15 and C11-C18 [10] (Fig. 1) that are crucial for the peptide's toxicity [16,17]. The same disulfide pattern and, generally, high sequence similarity is observed

in guanylin and uroguanylin (Fig. 1), the endogenous peptide hormones that physiologically target the same receptor, membrane guanylate cyclase C, which is located at the brush border of the surface of cells of the intestine, and STh is able to displace both of these hormones from their receptor binding site [5,18-22].

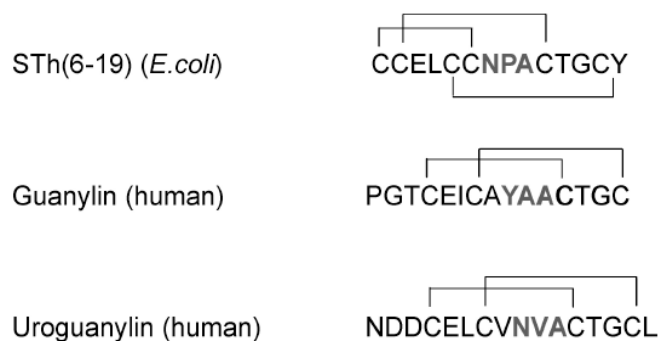


Fig. (1). Comparison of the amino acid sequences and disulfide bridge patterns of the heat stable enterotoxin from the human strain of enterotoxigenic *E. coli* and human hormones guanylin and uroguanylin. Possible binding region of the peptides are marked gray.

The initial step for the biological effect of STa is its binding to the extracellular domain of GC-C. This interaction leads to over activation of the intracellular GC-C cGMP kinase, which, in turn, results in an excessive signal to the cystic fibrosis transmembrane conductance regulator on the apical plasma membrane of small intestinal enterocytes that, in turn, elicits extreme chloride and fluid secretion [5,20,23,24].

Although the interaction of STa with GC-C is of crucial importance to this process, little is known about its molecu-

*Address correspondence to this author at the University of Bayreuth, Department of Biopolymers & Research Center for Bio-Macromolecules, D-95440 Bayreuth, Germany; E-mail: roesch@unibt.de

lar basis, and only a crystal structure (PDB: 1ETN, [25]) and an NMR solution structure of STI are known [26]. We are currently examining the structural basis of STh recognition by GC-C, and, as an initial step, we determined the solution structure of the chemically synthesized STh(6-19).

ENTEROTOXIN SYNTHESIS AND REFOLDING

The synthesis of STh(6-19) was performed using the Fmoc/But and maximal temporary protection strategy on a Syro II peptide synthesizer (MultiSynTech, Witten, Germany). The chemical procedure used 0.05 mmol of Fmoc-Tyr(tBu)-2-chlorotriptyl resin, an eightfold excess of each amino acid (Fmoc Cys(Trt)-OH, Fmoc-Gly-OH, Fmoc-Thr(tBu)-OH, Fmoc-Ala-OH), Fmoc-Pro-OH, Fmoc-Asn(Trt)-OH, Fmoc-Leu-OH and Fmoc-Glu(OtBu)-OH and 2-(1H-benzotriazole-1-yl)1,1,3,3-tetramethyluronium tetrafluoroborate/1-hydroxybenzotriazole (TBTU/HOBt) activation. Deprotection (2 h) and cleavage (100 mg peptide of resin) were achieved using 5 ml of a mixture of trifluoroacetic acid/thioanisole/ethanedithiole (90/8/2, vol/vol/vol). The acidic mixture was then precipitated three times with diethylether, dissolved in 10 % aqueous acetic acid and freeze dried. The crude toxin was purified by RP-HPLC on a C18 semi-preparative column (10 x 150 mm; Nucleosil) using a 40-min gradient of acetonitrile in 0.055% trifluoroacetic acid (10–80% B in 40 min, where B is 80 % acetonitrile/H₂O/0.05 % trifluoroacetic acid).

Oxidation of the reduced toxin was achieved by dissolving the purified peptide into 2 M acetic acid, and diluted to a peptide concentration of 0.015 mM in the presence of reduced/oxidized glutathione (molar ratio of peptide/GSH/GSSG was 1: 100: 10) and 2 M guanidine hydrochloride. The solution was adjusted to pH 8.0 with aqueous NH₄OH and stirred slowly at 4 °C for 7 d. The folding reaction was monitored by analytical HPLC. The solution was concentrated using a C18 SepPak (Waters) cartridge and finally lyophilized. Initial purification of the oxidized product was achieved by chromatography on a C8 column using the system above and yielding a purity of ~ 90 %. Finally, the product was highly purified on a C18 column using a 60-min gradient, resulting in a purity of 95 %. The quality of the product was confirmed by analytical HPLC, matrix-assisted laser desorption/ionization time of flight mass spectrometry (MALDI-MS) giving the correct mass in excellent agreement of the oxidized product. (M+H)_{calc} reduced: 1482.45; found: 1482.42; (M+H)_{calc} oxidized: 1476.41; found 1476.43.

NMR SPECTROSCOPY

Two-dimensional NMR spectra were recorded on Bruker DRX600 and AV800 spectrometers at 283 K with standard methods [27]. Standard ¹H-¹H correlated spectroscopy (COSY), ¹H-¹H total correlated spectroscopy (TOCSY) and ¹H-¹H homonuclear Overhauser enhancement spectroscopy (NOESY) were carried out with 4096x512 complex data points with excitation sculpting for water suppression [28] or coherence selection by pulsed field gradients [29]. Presaturation was applied for residual water suppression in experiments with the D₂O sample. ¹H-¹³C heteronuclear single quantum correlation (¹H-¹³C HSQC) and ¹H-¹³C-HMQC-TOCSY were used for ¹³C assignment and validation of the

¹H assignments. Peptide concentration was 3 mM, pH 3.0 in H₂O/D₂O (9: 1, v/v, 600μL) and in D₂O (99.98 %). For measurement in D₂O, STh was lyophilized repetitively from D₂O to exchange the amide protons and finally dissolved in D₂O, pH3. Spectra were processed and analyzed with in-house software and NMRView 5.2.2 [30].

STRUCTURE CALCULATIONS AND ANALYSIS

The total number of nontrivial unambiguous cross peaks in NOESY spectra was 190. The cross peaks were divided into three groups according to their relative intensities: strong with upper distance limit < 0.3 nm; medium, < 0.4 nm; and weak < 0.5 nm. Structure calculations were performed by using a modified *ab initio* SA protocol with the X-PLOR-NIH package [31]. The disulfide bonds were included explicitly. For each calculation 30 structures were calculated and 7 structures for each state were selected with the criteria for the lowest overall energy. Rasmol 2.7.3 [32,33] and PyMol [34] were used for molecular presentation. The geometry of the structures was analyzed using PROCHECK-NMR [35-37].

MD-SIMULATIONS

For further analysis and verification of our structural results we did an *ab initio* molecular dynamics simulation for STh and the hormones uroguanylin and guanylin [38]. The Amber 9 program package [39] and the *ff03* force-field [40,41] were used for the simulations of the three peptides. Each of them was constructed as an elongated peptide chain within the LEaP module of AMBER with the disulfide bonding as the only restraints.

The peptides were solvated in a TIP3P waterbox [42] with the dimensions of 80x60x40 Å, and for neutralization of the system sodium counterions were added. Calculations were performed at 286 K and an external pressure of 1 atm. At this conditions the systems were minimized and equilibrated using the program SANDER. Initially, the whole system was minimized for 1000 steps and the water molecules and the counterions were relaxed around the fixed solute with a 100-ps MD run. The systems were slowly heated stepwise to 286 K for equilibrating at each temperature. MD production runs of 2-ns duration were then performed for the systems. The MD data was analyzed by using the PTRAJ program. Root mean square deviation (r.m.s.d.) calculations of the heavy atoms were referenced to the NMR-structure of STh and the structures of the hormones deposited in the PDB (Guanylin: 1GUA, Uroguanylin; 1UYA), respectively.

RESULTS AND DISCUSSION

A detailed and well resolved solution structure of STh is needed for better understanding of processes that are involved into peptide recognition by its receptor. We thus chemically synthesized STh(6-19) and analyzed its NMR spectra. The chemical synthesis resulted in a peptide that was active in binding to the membrane proximal extracellular subdomain of human GCC with a nanomolar dissociation constant (Matecko *et al.*, unpublished).

The amide region of the proton NMR spectrum of STh showed the large dispersion of 2.5 ppm characteristic for a peptide with defined structure (Fig. 2). Using standard through-

involved in hydrogen bonds. During the iterative structure determination hydrogen bonds between N12 CO and C15 H^N; C18 H^N and C15 CO and C10 H^N and C7 CO were deduced.

For the structure calculation 190 experimentally derived distance restraints were obtained. Due to the observed line broadening by conformational exchange NOE peak intensities were classified very conservatively to include effects of dynamical averaging. The ten accepted structures out of ten calculated superimpose with a backbone r.m.s.d. of 0.89 Å and show only low violations of experimental and geometrical restraints (Table 2). A PROCHECK-NMR analysis of STh shows that 51 % of the residues of the accepted structures are found in the most favoured regions and an additional 49 % in the allowed regions of the Ramachandran plot (Table 2).

Table 2. Structural Statistics

NOE Statistics	
Total NOE number	190
Short range	21
Medium range	16
Long range	153
Deviation from Standard Geometry and Experimental Restraints	
Bonds	0.00095 ± 0.0002
Angles	0.169 ± 0.023
Distance restraints	0.0036 ± 0.0011
Ramachandrans plot statistics ^a	51 % / 31 % / 18 % / 0 %

^a Ramachandran plot statistics are determined by PROCHECK-NMR and are determined as follow: residues in most favored region, in additional region, in generously allowed region, and in disallowed region.

The solution structure of STh(6-19) is composed of an α -helical turn at its N-terminus region and two β -turns, between C11-C15, C15-C18 stabilised by the three disulfide bridges as mentioned before (Fig. 3).

The unavailability of the coordinates of the solution conformation from Garipey *et al.* [26] renders direct comparison of the structures impossible. Superimposing the present STh structure with the crystal structure (pdb code 1ETN [25]), however, resulted in a backbone r.m.s.d. of 1.6 Å for residues C7-C18, mainly due to different orientation of the carboxy terminus. Restricting the fit to residues C7-G16 lowers the rmsd to 0.9 Å, demonstrating similar conformations in solution and in the crystal. The receptor binding region of STh and the endogenic GC-C peptide ligands uroguanylin and guanylin is found to be from N12 - A14 for STh [44] and Y9 - A11 for guanylin [45]. In fact, these regions are highly solvent exposed for guanylin, uroguanylin, STp, and STh (Fig. 4).

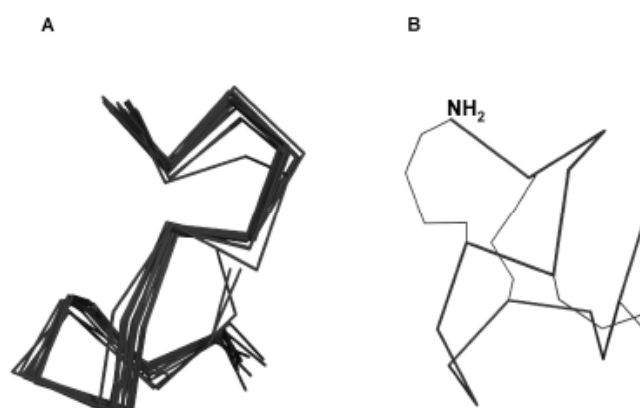


Fig. (3). (A) Overlay of 13 NMR derived structures of STh (6-19). (B) Presentation of disulfide bridges in STh (6-19).

Ab initio MD simulations with the NMR structures of guanylin, uroguanylin, and STh as starting structures show high flexibility of all three peptides in the loop regions (Fig.

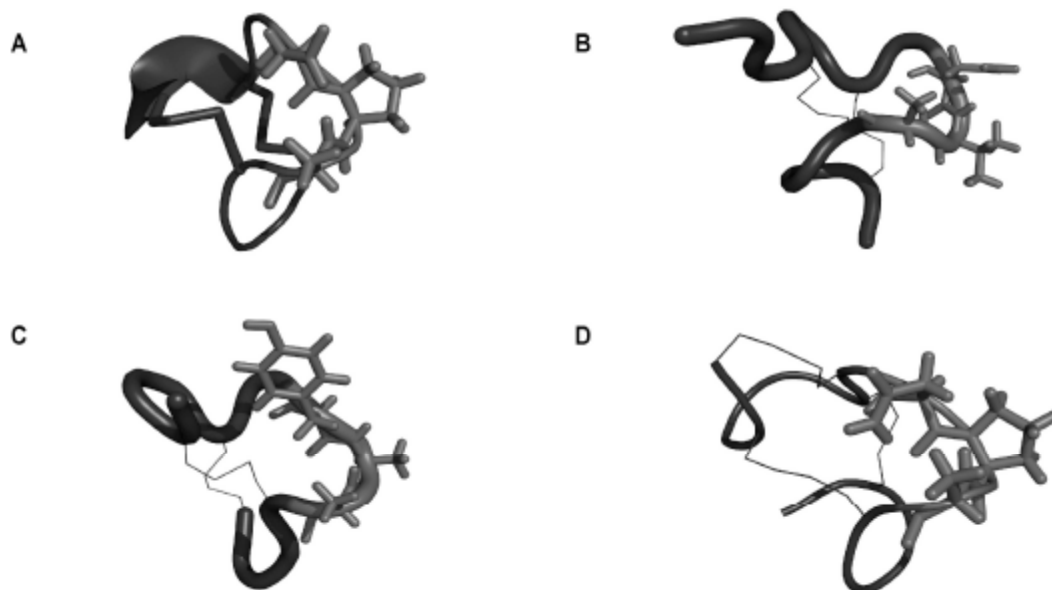


Fig. (4). Comparison of structures of (A) STp (5-17), PDB: 1ETN; (B) Uroguanylin human, PDB: 1UYA; (C) Guanylin human, PDB: 1GNA and (D) our calculated STh (6-19) structure. Possible binding sites are shown as sticks.

5). The r.m.s.d. values of heavy atoms are in the same range for all three peptides, despite the additional disulfide bridge in STh. To evaluate the importance of the three disulfide bridges for the structure calculation of STh, we performed the identical calculations as we did for the NMR-structure determination, however, without taking into account the disulfide bridges. The energetically most favourable 15 structures were virtually the same as from the calculation containing both, the NOE restraints and the disulfide bridges. The largest difference in the structures was in the N-terminal loop and the carboxy-terminus of the peptide which both were observed to show higher flexibility.

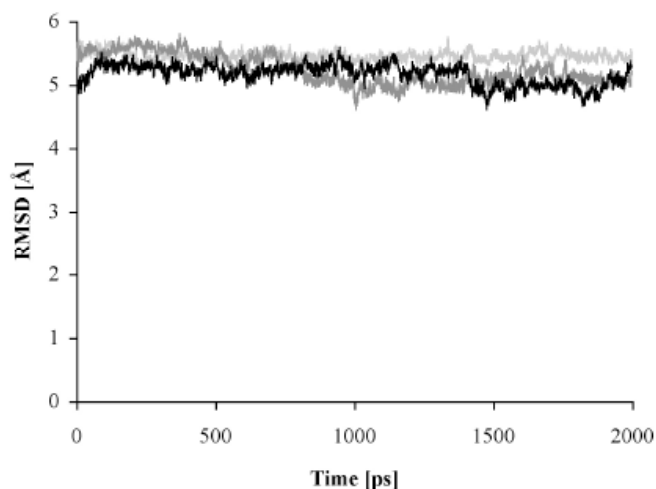


Fig. (5). RMSD of the heavy atoms during the MD. Sth (6-19) referenced to the present structure (black); Uroguanylin referenced to PDB: 1UYA (light grey); Guanylin referenced to PDB: 1GNA (dark grey).

In addition to differences in structure and dynamics, STh may act as a toxin because it does not contain the chymotrypsin cleavage site found in guanylin, the endogenous peptide that predominantly acts in the large intestine, as opposed to uroguanylin [45]. Chymotrypsin is an enzyme of the intestinal tract, and it cleaves after aromatic amino acids such as Y9 of guanylin (Fig. 1). In fact, the Y9NA10P double mutant of guanylin causes diarrhoea in suckling mice at much lower concentrations than the native peptide does [45]. Thus, one reason for the only transient action of native guanylin in the large intestine could be its rapid loss of its structure by enzymatic digestion.

ACKNOWLEDGEMENT

Financial support by the Deutsche Forschungsgemeinschaft to PR (MA2317/5-2) is gratefully acknowledged.

REFERENCES

- [1] Carter SR, Seiff SR. Cosmetic botulinum toxin injections. *Int Ophthalmol Clin* 1997; 37: 69-79.
- [2] Levine SR. Thrombolytic therapy for stroke: the new paradigm. *Hosp Pract (Minneapolis)* 1997; 15; 32(11): 57-64, 69-73.
- [3] Maseri A, Andreotti F. Targeting new thrombolytic regimens at specific patient groups: implications for research and cost-containment. *Eur Heart J* 1997; 18(Suppl F): F28-35.
- [4] Schmitt CK, Meysick KC, O'Brien AD. Bacterial toxins: friends or foes? *Emerg Infect Dis* 1999; 5: 224-34.
- [5] Hasegawa M, Shimonishi Y. Recognition and signal transduction mechanism of *Escherichia coli* heat-stable enterotoxin and its receptor, guanylate cyclase C. *J Pept Res* 2005; 65: 261-71.
- [6] Giblin MF, Gali H, Sieckman GL, et al. *In vitro* and *in vivo* evaluation of ¹¹¹In-labeled *E. coli* heat-stable enterotoxin analogs for specific targeting of human breast cancers. *Breast Cancer Res Treat* 2006; 98: 7-15.
- [7] Giblin MF, Sieckman GL, Shelton TD, Hoffman TJ, Forte LR, Volkert WA. *In vitro* and *in vivo* evaluation of ¹⁷⁷Lu- and ⁹⁰Y-labeled *E. coli* heat-stable enterotoxin for specific targeting of uroguanylin receptors on human colon cancers. *Nucl Med Biol* 2006; 33: 481-8.
- [8] Giblin MF, Sieckman GL, Watkinson LD, et al. Selective targeting of *E. coli* heat-stable enterotoxin analogs to human colon cancer cells. *Anticancer Res* 2006; 26: 3243-51.
- [9] Shimonishi Y, Hidaka Y, Koizumi M, et al. Mode of disulfide bond formation of a heat-stable enterotoxin (STh) produced by a human strain of enterotoxigenic *Escherichia coli*. *FEBS Lett* 1987; 215: 165-70.
- [10] Nair GB, Takeda Y. The heat-stable enterotoxins. *Microb Pathog* 1998; 24: 123-31.
- [11] Levine MM, Caplan ES, Waterman D, Cash RA, Hornick RB, Snyder MJ. Diarrhea caused by *Escherichia coli* that produce only heat-stable enterotoxin. *Infect Immun* 1977; 17: 78-82.
- [12] Moseley SL, Hardy JW, Hug MI, Echeverria P, Falkow S. Isolation and nucleotide sequence determination of a gene encoding a heat-stable enterotoxin of *Escherichia coli*. *Infect Immun* 1983; 39: 1167-74.
- [13] Lee CH, Moseley SL, Moon HW, Whipp SC, Gyles CL, So M. Characterization of the gene encoding heat-stable toxin II and preliminary molecular epidemiological studies of enterotoxigenic *Escherichia coli* heat-stable toxin II producers. *Infect Immun* 1983; 42: 264-8.
- [14] Rao MC. Toxins which activate guanylate cyclase: heat-stable enterotoxins. *Ciba Found Symp* 1985; 112: 74-93.
- [15] Lathe R, Hirth P, DeWilde M, Harford N, Lecocq JP. Cell-free synthesis of enterotoxin of *E. coli* from a cloned gene. *Nature* 1980; 284: 473-4.
- [16] Garipey J, Judd AK, Schoolnik GK. Importance of disulfide bridges in the structure and activity of *Escherichia coli* enterotoxin ST1b. *Proc Natl Acad Sci USA* 1987; 84: 8907-11.
- [17] Yoshimura S, Ikemura H, Watanabe H, et al. Essential structure for full enterotoxigenic activity of heat-stable enterotoxin produced by enterotoxigenic *Escherichia coli*. *FEBS Lett* 1985; 181(1): 138-42.
- [18] Hirayama T, Wada A, Hidaka Y, Fujisawa J, Takeda Y, Shimonishi Y. Expression of a truncated guanylate cyclase (GC-C), a receptor for heat-stable enterotoxin of enterotoxigenic *Escherichia coli*, and its dimer formation in COS-7 cells. *Microb Pathog* 1993; 15: 283-91.
- [19] Waldman SA, Kuno T, Kamisaki Y, et al. Intestinal receptor for heat-stable enterotoxin of *Escherichia coli* is tightly coupled to a novel form of particulate guanylate cyclase. *Infect Immun* 1986; 51: 320-6.
- [20] Hasegawa M, Hidaka Y, Matsumoto Y, Sanni T, Shimonishi Y. Determination of the binding site on the extracellular domain of guanylyl cyclase C to heat-stable enterotoxin. *J Biol Chem* 1999; 274: 31713-8.
- [21] Wada A, Hirayama T, Kitaura H, et al. Identification of ligand recognition sites in heat-stable enterotoxin receptor, membrane-associated guanylyl cyclase C by site-directed mutational analysis. *Infect Immun* 1996; 64: 5144-50.
- [22] Hasegawa M, Matsumoto-Ishikawa Y, Hijikata A, Hidaka Y, Go M, Shimonishi Y. Disulfide linkages and a three-dimensional structure model of the extracellular ligand-binding domain of guanylyl cyclase C. *Protein J* 2005; 24(5): 315-25.
- [23] Gardner P, Chao AC, De Sauvage F. STa receptors: Physiological and pathophysiological regulation of intestinal secretion by 5'-cyclic guanosine monophosphate. *Gastroenterology* 1995; 109(1): 325-7.
- [24] Harteneck C, Koesling D, Soling A, Schultz G, Bohme E. Expression of soluble guanylyl cyclase. Catalytic activity requires two enzyme subunits. *FEBS Lett* 1990; 272: 221-3.
- [25] Ozaki H, Sato T, Kubota H, Hata Y, Katsube Y, Shimonishi Y. Molecular structure of the toxin domain of heat-stable enterotoxin produced by a pathogenic strain of *Escherichia coli*. A putative binding site for a binding protein on rat intestinal epithelial cell membranes. *J Biol Chem* 1991; 266: 5934-41.

- [26] Garipey J, Lane A, Frayman F, *et al.* Structure of the toxic domain of the *Escherichia coli* heat-stable enterotoxin ST I. *Biochemistry* 1986; 25: 7854-66.
- [27] Cavanagh J, Fairbrother WJ, Palmer III AG, Skelton NJ. *Protein NMR Spectroscopy: Principles and Practice*. San Diego, CA: Academic Press; 2006.
- [28] Hwang TL, Shaka AJ. Water suppression that works. Excitation sculpting using arbitrary wave forms and pulsed gradients. *J Magn Res* 1995; 112A: 275-9.
- [29] Mori S, Abeygunawardana C, Johnson MO, van Zijl PC. Improved sensitivity of HSQC spectra of exchanging protons at short interscan delays using a new fast HSQC (FHSQC) detection scheme that avoids water saturation. *J Magn Res B* 1995; 108: 94-8.
- [30] Johnson BA. Using NMRView to visualize and analyze the NMR spectra of macromolecules. *Methods Mol Biol* 2004; 278: 313-52.
- [31] Schwieters CD, Kuszewski JJ, Tjandra N, Clore GM. The Xplor-NIH NMR molecular structure determination package. *J Magn Res* 2003; 160: 66-74.
- [32] Bernstein HJ. Recent changes to RasMol, recombining the variants. *Trends Biochem Sci* 2000; 25: 453-5.
- [33] Sayle R, Milner-White EJ. RASMOL: Biomolecular graphics for all. *Trends Biochem Sci* 1995; 20: 374-6.
- [34] DeLano WL. *The PyMOL Molecular Graphics System*. DeLano Scientific, Palo Alto, CA, USA. 2002.
- [35] Morris AL, MacArthur MW, Hutchinson EG, Thornton JM. Stereochemical quality of protein structure coordinates. *Proteins* 1992; 12: 345-64.
- [36] Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: A program to check the stereochemical quality of protein structures. *J Appl Cryst* 1993; 26: 283-91.
- [37] Laskowski RA, Rullmann JA, MacArthur MW, Kaptein R, Thornton JM. AQUA and PROCHECK-NMR: Programs for checking the quality of protein structures solved by NMR. *J Biomol NMR* 1996; 8: 477-86.
- [38] Simmerling C, Strockbine B, Roitberg AE. All-Atom Structure Prediction and Folding Simulations of a Stable Protein. *J Am Chem Soc* 2002; 38: 11258-9.
- [39] Case DA, Darden TA, Cheatham TE, *et al.* AMBER 9, University of California, San Francisco, 2006.
- [40] Duan Y, Wu C, Chowdhury S, *et al.* A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *J Comput Chem* 2003; 24: 1999-2012.
- [41] Lee MC, Duan Y. Distinguish protein decoys by using a scoring function based on a new AMBER force field, short molecular dynamics simulations, and the generalized born solvent model. *Proteins: Struct Funct Bioinf* 2004; 55: 620-34.
- [42] Price DJ, Brooks III CL. A modified TIP3P water potential for simulation with Ewald summation. *J Chem Phys* 2004; 121: 10096-103.
- [43] Sharma D, Rajarathnam K. ¹³C NMR chemical shifts can predict disulfide bond formation. *J Biomol NMR* 2000; 18: 165-7.
- [44] Carpick BW, Garipey J. Structural characterization of functionally important regions of the *Escherichia coli* heat-stable enterotoxin STIb. *Biochemistry* 1991; 30: 4803-9.
- [45] Carpick BW, Garipey J. The *Escherichia coli* heat-stable enterotoxin is a long-lived superagonist of guanylin. *Infect Immun* 1993; 61: 4710-5.

Received: August 15, 2008

Revised: October 27, 2008

Accepted: November 7, 2008

© Matecko *et al.*; Licensee Bentham Open

This is an open access article licensed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted, non-commercial use, distribution and reproduction in any medium, provided the work is properly cited.

Einzelarbeit C

Claudia Knake, Björn M. Burmann, Kristian Schweimer, Irena Matečko und Paul Rösch:

The Membrane Proximal Extracellular Domain of Human hGC-B Folds Independently.
Journal of Biomolecular Structure & Dynamics **26**, 465 – 472 (2009)

The Membrane Proximal Extracellular Domain of Human hGC-B Folds Independently

<http://www.jbsdonline.com>

Claudia Knake
Björn M. Burmann
Kristian Schweimer
Irena Matečko
Paul Rösch*

Department of Biopolymers and Research
Center for Bio-Macromolecules
University of Bayreuth
Universitätsstraße 30
95440 Bayreuth, Germany

Abstract

Human Guanylyl Cyclase B (hGC-B) is a single-transmembrane receptor protein which upon binding C-type natriuretic peptide (CNP) to its extracellular domain catalyzes the intracellular conversion of GTP to the second messenger cGMP. cGMP in turn affects various physiological processes such as smooth muscle contraction, cell proliferation, phototransduction, and salt as well as fluid homeostasis. The 3-dimensional binding site of the peptide hormone is unknown, and the binding mechanism is not yet understood. Therefore, a model of the C-terminal moiety of the extracellular domain of human GC-B containing the potential binding site was derived from the crystal structure of (GC-A). The selected protein sequence was provided with an N-terminal TEV-cleavage site and fused with a 109 aa thioredoxin-tag and a hexahistidine-tag. The identity of the purified 25 kDa protein was confirmed by protein mass fingerprint and its secondary structure was determined by CD- and NMR-spectroscopy. The protein proved to be properly folded with the observed secondary structure matching the predicted secondary structure and the homologous structure in the extracellular domain of GC-A. Size exclusion chromatography confirmed the monomeric state of P-hGC-B.

Key words: Guanylate cyclase; Natriuretic peptide; cGMP; ANP; GCB; Peptide hormone; Receptor; Transmembrane; Extracellular domain; Second messenger.

Introduction

Human Guanylyl Cyclase B (hGC-B) is one of seven particulate guanylyl cyclases that have been identified in mammals so far. This receptor family bears resemblance to soluble guanylyl cyclases and adenylyl cyclases and belongs to a larger family of single-transmembrane receptors that also includes growth hormone and cytokine receptors. In general, the structure of a typical GC receptor consists of an extracellular N-terminal domain (ECD), a transmembrane moiety (TM) and an intracellular part comprising a protein kinase-homology domain (KHD) and the cyclase catalytic region (GCD). On ligand binding to the ECD, the particulate GCs undergo a conformational change enabling them to convert GTP to cGMP at their intracellular domain thus participating in transmembrane signal transduction. This signaling pathway is implicated in a variety of physiological processes such as smooth muscle contraction, cell proliferation, phototransduction, and salt as well as fluid homeostasis (1). In humans, the ligands of three particulate GCs (hGC-A, hGC-B, hGC-C) have been identified. hGC-A, which regulates blood pressure by diuretic and natriuretic effects on the kidney, binds atrial natriuretic peptide (ANP) and B-type natriuretic peptide (BNP) (2, 3); hGC-B, which exhibits mainly vasorelaxant effects but is also important for bone growth (4) and the attenuation of vascular inflammation (5), is activated by C-type natriuretic peptide (CNP). ANP, BNP as well as CNP bind to the natriuretic clearance receptor (NPR-C), which lacks most of the intracellular domains found in the other receptors and does not possess guanylyl cyclase activity. hGC-C is mainly expressed in the intestine and

*Phone: +49921553540
Fax: +4992116490459
Email: roesch@uni-bt.de

is thought to regulate the fluid and electrolyte balance. hGC-C binds guanylin, uroguanylin, and the *E. coli* heat stable enterotoxin STA. Specific ligands for the other four known GCs – GC-D to GC-G – have not been identified so far.

As several cardiovascular and skeletal disorders are associated with dysfunctions in natriuretic peptide signaling, this pathway holds promise as a therapeutic target. Defects in the hGC-B gene (*NPR2*) on chromosome 9, for example, are responsible for a form of human dwarfism called *acromesomelic dysplasia type Maroteaux (AMDM)* (6) that is characterized by deformity of forearms, forelegs, fingers and toes, spine and head abnormalities, and delayed motor development.

The functional unit of GC-A and -B is a homodimer, and even without ligand the monomers undergo spontaneous self-association to form dimers or oligomers (7). The ECDs of both receptors contain six cysteine residues, which form three intramolecular disulfide bonds with a 1-2, 3-4, 5-6 pairing pattern. Both ECDs possess multiple N-glycosylation sites and partially as well as fully glycosylated forms can be detected. The functional and structural importance of glycosylation is controversial, and glycosylation may or may not be required for proper folding, targeting, and ligand binding (8-10).

Most of the work to characterize the natriuretic peptide receptors was performed on GC-A, which exhibits a high degree of sequence similarity to hGC-B. The sequence similarity between the two ECDs is approximately 44% and it is even higher comparing the intracellular domains (KHD: 63%; GCD: 88%) (11). Therefore, structure and ligand binding mechanism are believed to be similar as well, and the structural information gained on GC-A should well serve as a model for understanding hGC-B. Accordingly, a model of the C-terminal membrane proximal moiety of the ECD of hGC-B containing part of the ligand binding site was derived from the crystal structure of GC-A (12) and was named P-hGC-B. The purified protein is monomeric and properly folded according to size exclusion chromatography, CD- and NMR-data, with the observed secondary structure matching the predicted secondary structure as well as the homologous structure in the extracellular domain of GC-A.

Material and Methods

Homology Modeling

PSI-BLAST (13) was used to find protein sequences similar to the ECD of hGC-B, and significant homology was found to the ligand binding regions of the ANP receptor family. The sequences of the ECDs of GC-A and hGC-B were aligned using CLUSTALW (14). The sequence similarity between both extracellular domains is approximately 44%, which renders the crystal structure of hGC-A a suitable template for comparative modeling with Modeller9v2 (15). Template coordinates were taken from the GC-A ECD crystal structure (PDB code: 1DP4), and the sequence alignment of the ECDs of hGC-B and GC-A was based on the results from 3D-PSSM (16). A structural model for the hGC-B EDC was calculated for sequence positions 23-458. Using the model of the extracellular domain of hGC-B the structure of P-hGC-B was modeled with the same programs. A disulfide bond was introduced using the AMBER program and molecular stability was checked using the same package.

Molecular Dynamics (MD) Simulations

The Amber 9 program package (17) and the ff03 force-field (18) were used for the simulation of P-hGC-B. It was solvated in a TIP3P water box (19) of 80×50×60 Å³, and sodium counter ions were added *in silicio* for neutralization of the system using the LEAP program. Additionally, the disulfide bond between Cys208 and

Cys217 was introduced as a calculational restraint. Calculations were performed at 298 K and an external pressure of 1 atm. At these conditions the system was minimized and equilibrated using the program SANDER. Initially, the whole system was minimized for 1000 steps, and the water molecules and counterions were relaxed around the fixed solute with a 100-ps MD run. MD production runs of 5-ns duration were then performed for the P-hGC-B. The MD data was analyzed using the PTRAJ program.

Expression and Purification

A 720 bp synthetic gene adapted to *E. coli* codon usage harboring the sequence the P-hGC-B was obtained in the cloning vector pCR2.1-PGCB (Eurofins Medigenomix, Martinsried) and subcloned into the *E. coli* expression vector pET-32a(+) (20) (Novagen) using SacI/HindIII cloning sites. The construct expressed a fusion protein containing a N-terminal thioredoxin and 6His-tag followed by a TEV cleavage site and P-hGC-B. The resulting pET-32a(+)-TEV-PGC-B vector was transformed into *E. coli* Rosetta-gami B(DE3) (Novagen). After transformation the cells were cultured on LB agar plates containing ampicillin (100 µg/ml), kanamycin (15 µg/ml), tetracyclin (12.5 µg/ml), and chloramphenicol (34 µg/ml) at 37 °C overnight. For an induction culture, the cells were grown at 37 °C in M9 medium containing 2 mM MgSO₄, 10 µM Fe(III)-citrate, 0.1 mM CaCl₂, trace element solution TS2 (21), MEM vitamin solution (Gibco, Invitrogen), antibiotics and 1.5 g/l (NH₄)₂SO₄ and 4 g/l glucose as sole nitrogen and glucose source until an OD₆₀₀ of 0.6 was reached. The cell culture was cooled down to 25 °C and expression of the fusion protein was induced with 100 µM IPTG at an OD₆₀₀ of 0.7-0.8. Cells were harvested 3 hours after induction and stored at -20 °C. To obtain ¹⁵N labeled protein, (NH₄)₂SO₄ in the M9 medium was replaced by (¹⁵NH₄)₂SO₄. The frozen cell pellet was resuspended in lysis buffer containing 50 mM sodium phosphate (pH 7.8), 500 mM NaCl, lysozyme, DNase and protease inhibitor cocktail (Roche), and was lysed with a Microfluidizer. The lysate was fractionated by centrifugation (6000 rpm, 20 min). The supernatant containing the soluble His6-tagged protein was applied to a 5 ml HiTrap Chelating column (GE Healthcare, Munich, Germany) charged with Co²⁺ and pre-equilibrated with binding buffer (50 mM sodium phosphate, pH 7.8, 500 mM NaCl, 10 mM imidazole). After washing the column with three column volumes binding buffer, the protein was eluted using a stepwise gradient including 5, 20, 100% elution buffer (50 mM sodium phosphate, pH 7.8, 500 mM NaCl, 500 mM imidazole). Protein was eluted with 108 mM imidazole. The whole purification protocol was performed on an Äktapurifier (GE Healthcare, Germany). The fractions showing protein peaks were analyzed on an SDS-gel. Fractions containing the desired protein were pooled. TEV-cleavage was performed during dialysis against phosphate buffer (50 mM sodium phosphate, 500 mM NaCl; TEV) overnight at 4 °C. Successful cleavage was confirmed by SDS-PAGE. TEV protease and cleaved Trx-6His-tag were separated from P-hGC-B by Co²⁺ affinity chromatography. After purification, the protein was dialyzed against 20 mM sodium phosphate buffer (pH 7.0) and 100 mM NaCl at 4 °C for at least 4 hours. The dialyzed protein was concentrated using VIVAspin columns (Satorius, Göttingen, Germany) and stored at 4 °C. The identity of the purified protein was confirmed by protein mass fingerprinting (ZMMK Köln).

Circular Dichroism Spectroscopy (CD)

P-hGC-B samples (2.5 µM) were diluted in 20 mM sodium phosphate buffer pH 7.0. Spectra were recorded at 25 °C from 190 nm-260 nm with 50 nm/min scanning speed (J-810 S spectropolarimeter, CDF-426S temperature control unit, JASCO International, Tokyo, Japan) in a 1 mm path length quartz cuvette (Hellma, Müllheim, Germany). 6 spectra were accumulated and corrected for buffer effects, the CD-data was processed with the CDSSTR algorithm at the Dichroweb server (22).



Figure 1: P-hGC-B construct. **(A)** Schematic representation of hGC-B monomer. ECD, extracellular domain; D-hGC-B, membrane-distal part of the ECD; P-hGC-B, membrane-proximal part of the ECD; TM, trans-membrane region; KHD, kinase homology domain; GCD, guanylyl cyclase domain. **(B)** P-hGC-B amino acid sequence. APGD was used as linker-sequence (green). C83S and the intra-chain disulfide bond are marked red **(C)** Model of the ECD P-hGC-B calculated with Modeller9v2. Green: amino acids on the incision sites. **(D)** Model structure of P-hGC-B based on the model structure of ECD P-hGC-B. Green: linker peptide.

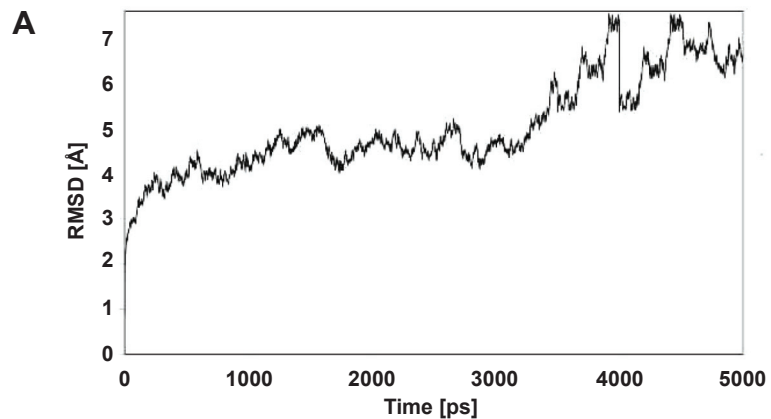
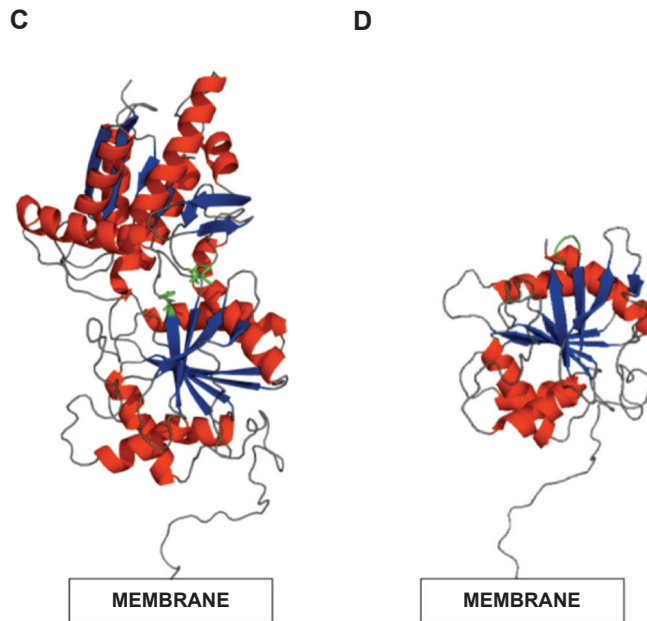
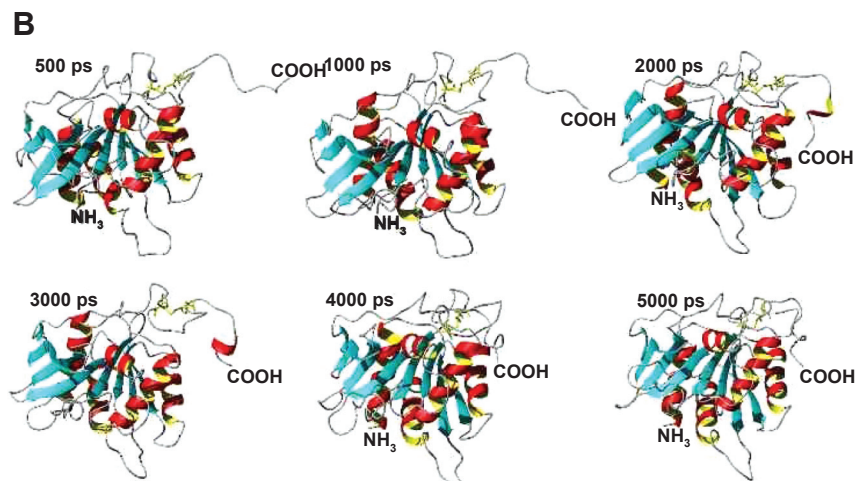


Figure 2: **(A)** Rmsd-values of the atomic coordinates for P-hGC-B as a function of the simulation time referenced to the first calculated structure. **(B)** Snapshots taken from the MD-simulation of P-hGC-B showing a high degree of stability in the center through the formation of several β -strands (turquoise). The structure is further stabilized by several α -helices (red) surrounding the core of β -strands.



Size Exclusion Chromatography

Size exclusion chromatography was carried out using two coupled 1 × 30 cm Superdex 75-HR columns (GE Healthcare, Germany) on an Äktapurifier at room temperature. P-hGC-B (100 μM) was eluted at a flow rate of 0.5 ml/min with 20 mM sodium phosphate buffer and 100 mM NaCl at pH 6.9. For calibration, cytochrome *c*, ovalbumin, chymotrypsinogen a, and albumin were used.

NMR Spectroscopy

The sample contained ¹⁵N P-hGC-B, 500 μM, in 20 mM sodium phosphate buffer pH 7.0 with 100 mM NaCl, and the NMR measurement was performed at 298 K with an 800 MHz Bruker AV800 spectrometer using standard techniques for recording and water suppression. NMR data was analyzed with NMRView 5.2.2 (23) and in house routines.

Results and Discussion

Modeling and Stability

The ECD of hGC-B as well as the corresponding domains of hGC-C and hGC-A belong to the class of periplasmic proteins with type 1 periplasmic binding protein fold (<http://scop.mrc-lmb.cam.ac.uk/scop>). This type of fold has a two domain architecture, each with a six stranded parallel β-sheet bordered with α-helices and a flexible linker region connecting these two domains. The ECD exhibits an N-terminal membrane distal domain and a C-terminal membrane proximal domain, which is assumed to be a ligand binding domain (24–27). From the model of hGC-B EDC (Fig. 1) the membrane proximal domain containing the ligand binding site could be derived by cutting the sequence at three sites, before Tyr276 and Thr368 in an antiparallel β-strand region of the proximal domain and before Ser123 in a helix region. Tyrosin and threonin were chosen as a section site since they are common in β-strands and the distance between them is approximately 5 Å, so they could easily be reconnected with the loop to yield a single continuous strand. This construct leaves the putative ligand binding site intact. The decision to start the domain with Ser123-Ala124-Pro125 before helix 1 was based on the stabilizing effect of those amino acids to the helix 1 (28). Cys205Ser was introduced to avoid the formation of unspecific inter- and intramolecular disulfide bonds. The whole system was stable in AMBER MD calculations after approximately 300 ns (Fig. 2A) as evidenced by the merely slight increase in the rmsd-value with ongoing simulation time. Snapshots taken from the simulation (Fig. 2B) show that the early stabilization of the structure is mainly due to the formation of the central β-strands, followed by better definition of the surrounding α-helices. Flexibility of the C-terminus leads to an increase in the rmsd-value at simulation times from 3.5 to 5.0 ns (Fig. 2). The snapshot at 4 ns (Fig. 2B), in contrast to all other snapshots, shows the C-terminus pointing towards the rear of the figure, resulting in a 90° deviation from its prevailing position. This flexibility, however, seems not to affect the overall fold and stability of P-hGC-B.

Expression and Purification

From different temperatures (20 °C, 25 °C, 37 °C) and IPTG-concentrations (100 μM, 200 μM, 500 μM, 1 mM) an optimal choice of parameters turned out to be 25 °C with 100 μM IPTG for induction (Fig. 3A). The protein mostly accumulated in the supernatant after cell lysis (Fig. 3B), whereas at higher temperatures or higher IPTG-concentrations it was mainly insoluble because of inclusion body formation. The cleavable N-terminal Trx-6His-tag allowed purification of soluble P-hGC-B by Co²⁺ affinity chromatography (Fig. 3C). The protein fractions were pooled and TEV cleavage of the N-terminal Trx-6His-tag was performed during dialysis against lysis

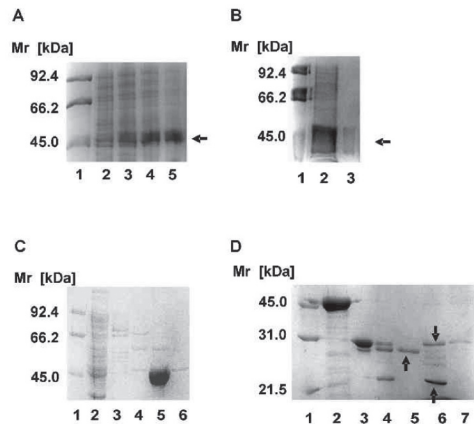


Figure 3: Expression and purification of soluble P-hGC-B. (A) Induction in LB: Rosetta-gami B(DE3); 100 μ M IPTG; 25 $^{\circ}$ C; lane 1: molecular weight standard (MW); lane 2-5: before induction, 1h, 2h, 3h after induction; arrow indicates P-hGC-B. (B) Solubility of P-hGC-B after cell lysis; lane 1: MW; lane 2: supernatant; lane 3: pellet; arrow indicates P-hGC-B. (C) First purification step. Protein fractions after first purification on the Äktapurifier using a cobalt column; lane 1: MW, lane 2: flow-through, lane 3: wash with binding buffer (10 mM imidazole); lane 4: elution with 34.5 mM imidazole; lane 5: elution with 108 mM imidazole; lane 6: elution with 500 mM imidazole. (D) Second purification step. Lane 1: MW; lane 2: P-hGC-B after first cobalt column; lane 3: TEV protease; lane 4: P-hGC-B after dialysis with TEV overnight; lane 5: P-hGC-B after second purification on the cobalt column (flow-through); arrow indicates purified P-hGC-B; lane 6: elution with 34.5 mM imidazole; upper arrow indicates TEV protease, lower arrow indicates cut Trx-6His-tag; lane 7: elution with 500 mM imidazole.

buffer overnight (Fig. 3D). TEV protease and cut Trx-tag each containing a 6His-tag were removed from P-hGC-B by Co^{2+} affinity chromatography (Fig. 3D). Purity and correct amino acid sequence were confirmed by peptide mass fingerprint.

Structural Analysis

The far UV CD-spectrum of P-hGC-B exhibits characteristic features of α -helical and β -sheet structure elements (Fig. 4A). The estimated content of secondary structures was 29% α -helix and 23% β -strand, and the PredictProtein program (29) predicted 26% α -helix and 22% β -strand, in nearly perfect agreement with the experimentally derived data.

The high dispersion of the signals in the ^1H - ^{15}N HSQC of ^{15}N labeled P-hGC-B (Fig. 4B) indicates a well defined tertiary structure of the P-hGC-B protein.

Oligomerization State

In the absence of ligand, GC-A and GC-B have been observed to exist as homomers (7), and purified ECD of GC-A tends to form dimers (30). From crystallization data of dimerized glycosylated GC-A ECD, dimerization was proposed to occur *tail-to-tail* via the membrane proximal subdomain (12), from crystallized NPR-C ECD, however, *head-to-head* dimerization via membrane distal domain interaction was inferred (31). In the case of P-hGC-B, size exclusion chromatography shows the isolated model membrane proximal domain to exist as a monomer (Fig. 4C), disfavoring the *tail-to-tail* dimerization model under the current experimental conditions.

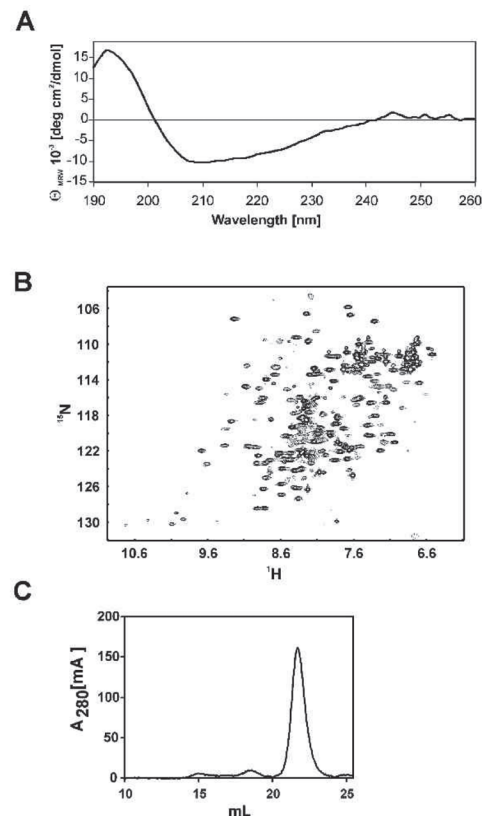


Figure 4: Structure analysis of recombinant P-hGC-B. (A) Far UV spectrum of 2.5 μ M P-hGC-B in 20 mM sodium phosphate buffer. Spectrum is typical for proteins containing α -helices and β -sheets. (B) ^1H - ^{15}N HSQC spectrum of P-hGC-B (500 μ M) in 20 mM sodium phosphate pH 7.0, 100 mM NaCl, 1 mM NaN_3 , at 298 K; 800 MHz. (C) Size-exclusion chromatography of P-hGC-B. Protein concentration was 100 μ M in 20 mM sodium phosphate buffer and 100 mM NaCl. The peak at an elution volume of 21.12 ml indicates monomeric P-hGC-B.

Conclusions

We here demonstrated the successful design and expression of P-hGC-B, a model for the membrane proximal, ligand binding ECD of hGC-B. The isolated domain forms a stable tertiary structure independent of the rest of the protein, and this

domain is not sufficient for protein oligomerization. In addition, P-hGC-B results in very well resolved HSQC-spectra. Thus, the present study may serve as a starting point to determine the ligand binding activities and, eventually, paving the way to determine the tertiary structure of the receptor:ligand complex by NMR-spectroscopy and X-ray crystallography.

Ligand binding of P-hGC-B, however, may be compromised since the membrane-distal subdomain of the ECD has been truncated and, although this subdomain seems to be mainly involved in dimerization (31-33), its contribution to ligand binding is clearly evidenced in GC-A (34, 35). Therefore, the ligand binding properties of P-hGC-B are currently studied by fluorescence spectroscopy and NMR HSQC-titration in order to define more clearly the role of the various parts of the extracellular domain to ligand binding, dimerization, and signal transduction.

Acknowledgments

This work was supported by the Elite-Network of Bavaria. Financial support by the Deutsche Forschungsgemeinschaft to PR (grant MA2317/5-2) is gratefully acknowledged.

References and Footnotes

1. K. A. Lucas, G. M. Pitari, S. Kazerounian, I. Ruiz-Stewart, J. Park, S. Schulz, K. P. Chepenik, S. A. Waldman. *Pharmacol Rev* 52, 375-414 (2000).
2. S. Suga, K. Nakao, K. Hosoda, M. Mukoyama, Y. Ogawa, G. Shirakami, H. Arai, Y. Saito, Y. Kambayashi, K. Inouye. *Endocrinology* 130, 229-239 (1992).
3. K. J. Koller, D. G. Lowe, G. L. Bennett, N. Minamino, K. Kangawa, H. Matsuo, D. V. Goeddel. *Science* 252, 120-123 (1991).
4. A. Yasoda, Y. Ogawa, M. Suda, N. Tamura, K. Mori, Y. Sakuma, H. Chusho, K. Shiota, K. Tanaka, K. Nakao. *J Biol Chem* 273, 11695-11700 (1998).
5. A. K. Kierner, M. D. Lehner, T. Hartung, A. M. Vollmar. *Endocrinology* 143, 846-852 (2002).
6. C. F. Bartels, H. Bukulmez, P. Padayatti, D. K. Rhee, C. van Ravenswaaij-Arts, R. M. Pauli, S. Mundlos, D. Chitayat, L. Y. Shih, L. I. Al-Gazali, S. Kant, T. Cole, J. Morton, V. Cormier-Daire, L. Faivre, M. Lees, J. Kirk, G. R. Mortier, J. Leroy, B. Zabel, C. A. Kim, Y. Crow, N. E. Braverman, F. van den Akker, M. L. Warman. *Am J Hum Genet* 75, 27-34 (2004).
7. M. Chinkers, E. M. Wilson. *J Biol Chem* 267, 18589-18597 (1992).
8. R. Fenrick, N. Bouchard, N. McNicoll, A. De Lean. *Mol Cell Biochem* 173, 25-32, (1997).
9. A. Nandi, R. Mathew, S. S. Visweswariah. *Protein Expr Purif* 8, 151-159, (1996).
10. Y. Hidaka, Y. Matsumoto, Y. Shimonishi. *FEBS Lett* 526, 58-62 (2002).
11. K. J. Koller, D. V. Goeddel. *Circulation* 86, 1081-1088 (1992).
12. F. van den Akker, X. Zhang, M. Miyagi, X. Huo, K. S. Misono, V. C. Yee. *Nature* 406, 101-104 (2000).
13. S. F. Altschul, T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, D. J. Lipman. *Nucleic Acids Res* 25, 3389-3402 (1997).
14. R. Chenna, H. Sugawara, T. Koike, R. Lopez, T. J. Gibson, D. G. Higgins, J. D. Thompson. *Nucleic Acids Res* 31, 3497-3500 (2003).
15. A. Sali, T. L. Blundell. *J Mol Biol* 234, 779-815 (1993).
16. L. A. Kelley, R. M. MacCallum, M. J. Sternberg. *J Mol Biol* 299, 499-520 (2000).
17. D. A. Case, T. A. Darden, T. E. Cheatham, C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, K. M. Merz, D. A. Pearlman, M. Crowley, R. C. Walker, W. Zhang, B. Wang, S. Hayik, A. Roitberg, G. Seabra, K. F. Wong, F. Paesani, X. Wu, S. Brozell, V. Tsui, H. Gohlke, L. Yang, C. Tan, J. Mongan, V. Hornak, G. Cui, P. Beroza, D. H. Mathews, C. Schafmeister, W. S. Ross, P. A. Kollman, Amber 9 (2006).
18. M. C. Lee, Y. Duan. *Proteins: Struct, Funct, Bioinf* 55, 620-634 (2004).
19. D. J. Price, C. L. Brooks III. *J Chem Phys* 121, 10096-10103 (2004).
20. E. R. LaVallie, E. A. DiBlasio, S. Kovacic, K. L. Grant, P. F. Schendel, J. M. McCoy. *Biotechnology* 11, 187-193 (1993).
21. O. Meyer, H. G. Schlegel. *Annu Rev Microbiol* 37, 277-310 (1983).
22. L. Whitmore, B. A. Wallace. *Nucleic Acids Res* 32, W668-673 (2004).
23. B. A. Johnson, R. A. Blevins. *J Biomol NMR* 4, 603-614 (1994).
24. M. Hasegawa, Y. Shimonishi. *J Pept Res* 65, 261-271 (2005).
25. C. Harteneck, D. Koesling, A. Soling, G. Schultz, E. Bohme. *FEBS Letters* 272, 221-223 (1990).
26. M. Hasegawa, Y. Hidaka, Y. Matsumoto, T. Sanni, Y. Shimonishi. *J Biol Chem* 274, 31713-31718 (1999).
27. Y. Hidaka, Y. Matsumoto, Y. Shimonishi. *FEBS Letters* 526, 58-62 (2002).
28. J. S. Richardson, D. C. Richardson. *Science* 240, 1648-1652 (1988).

29. B. Rost, G. Yachdav, J. Liu. *Nucleic Acids Res* 32, W321-326 (2004).
30. K. S. Misono, N. Sivasubramanian, K. Berkner, X. Zhang. *Biochemistry* 38, 516-523 (1999).
31. X. He, D. Chow, M. M. Martick, K. C. Garcia. *Science* 293, 1657-1662 (2001).
32. K. Misono, H. Ogawa, Y. Qiu, C. M. Ogata. *Peptides* 26, 957-968 (2005).
33. X. He, A. Dukupati, K. C. Garcia. *J Mol Biol* 361, 698-714 (2006).
34. H. Ogawa, Y. Qiu, C. M. Ogata, K. S. Misono. *J Biol Chem* 279, 28625-28631 (2004).
35. F. v. d. Akker. *J Mol Biol* 311, 923-937 (2001).

Date Received: May 5, 2008

Communicated by the Editor Ramaswamy H. Sarma

Einzelarbeit D

Björn M. Burmann, Xiao Luo, Paul Rösch, Markus C. Wahl und Max E. Gottesman:

Fine tuning of the *E. coli* NusB:NusE complex affinity to RNA is required for processive antitermination. *Nucleic Acids Research* **38**, 314 – 326 (2010)

Fine tuning of the *E. coli* NusB:NusE complex affinity to *BoxA* RNA is required for processive antitermination

Björn M. Burmann¹, Xiao Luo^{2,3}, Paul Rösch^{1,*}, Markus C. Wahl^{2,3} and Max E. Gottesman⁴

¹Lehrstuhl Biopolymere und Forschungszentrum für Bio-Makromoleküle, Universität Bayreuth, Universitätsstraße 30, 95447 Bayreuth, ²Max-Planck-Institut für biophysikalische Chemie, Makromolekulare Röntgenkristallographie, Am Faßberg 11, 37077 Göttingen, ³Freie Universität Berlin, Fachbereich Biologie-Chemie-Pharmazie, Institut für Chemie und Biochemie, AG Strukturbiochemie, Takustr. 6, 14195 Berlin, Germany and ⁴Department of Microbiology and Institute of Cancer Research, Columbia University Medical Center, New York, NY 10032, USA

Received July 22, 2009; Revised and Accepted August 20, 2009

ABSTRACT

Phage λ propagation in *Escherichia coli* host cells requires transcription antitermination on the λ chromosome mediated by λ N protein and four host Nus factors, NusA, B, E (ribosomal S10) and G. Interaction of *E. coli* NusB:NusE heterodimer with the single stranded *BoxA* motif of λ *nutL* or λ *nutR* RNA is crucial for this reaction. Similarly, binding of NusB:NusE to a *BoxA* motif is essential to suppress transcription termination in the ribosomal RNA (*rrn*) operons. We used fluorescence anisotropy to measure the binding properties of NusB and of NusB:NusE heterodimer to *BoxA*-containing RNAs differing in length and sequence. Our results demonstrate that *BoxA* is necessary and sufficient for binding. We also studied the gain-of-function D118N NusB mutant that allows λ growth in *nusA1* or *nusE71* mutants. *In vivo* λ burst-size determinations, CD thermal unfolding measurements and X-ray crystallography of this as well as various other NusB D118 mutants showed the importance of size and polarity of amino acid 118 for RNA binding and other interactions. Our work suggests that the affinity of the NusB:NusE complex to *BoxA* RNA is precisely tuned to maximize control of transcription termination.

INTRODUCTION

Phage λ -mediated antitermination in *Escherichia coli* enables RNA polymerase (RNAP) to read through early transcription termination sites on the phage chromosome (1). Antitermination is regulated *via* the direct interaction of N protein and the transcription elongation complex

(TEC) formed by RNA, RNAP and the Nus (N utilization substance) host factors NusA, NusB, NusE (S10 ribosomal protein) and NusG (2–4). N-mediated antitermination is coupled to transcription of the phage λ *nut* RNA sites, each consisting of the single stranded *BoxA* and the palindromic stem-loop *BoxB* linked by a *spacer* sequence to which NusA binds (5). N interacts with *BoxB* and converts the TEC to a termination-resistant form (6,7). Binding of λ N to *BoxB* results in an indirect interaction with RNAP through NusA (8,9). NusB interacts with the *nut* site by binding to *BoxA*, an interaction that is \sim 10-fold strengthened upon NusE:NusB heterodimer formation (10–13). The NusB:NusE:RNA ternary complex is proposed to associate with RNAP through NusE (1,14,15). A similar complex in which λ N is replaced by phage HK022 Nun protein induces transcription arrest on the λ chromosome (7).

In addition to its involvement in transcription, NusE participates in translation as part of the 30S ribosomal subunit (16–18).

A termination-resistant TEC also assembles during transcription of *rrn* operons in *E. coli* and other bacteria (19,20). In addition to Nus factors, ribosomal proteins S4, S2, L4 and L13 participate in transcription regulation (21,22). *BoxA* is highly conserved in all seven *E. coli* *rrn* operons. A promoter-proximal *BoxB*-like element is present but is not required for *rrn* antitermination (23). As is the case with λ , formation of the ternary NusB:NusE:*BoxA* complex is a key step during *rrn* processive antitermination (13).

The structure of the NusB:NusE ^{Δ loop} complex, in which the 22 residue ribosome-binding loop of NusE was deleted, has recently been determined (24). Analysis of this structure and other data (25) suggest that NusE is the active partner of the complex and that NusB mainly acts as a loading factor for NusE, a notion that is

*To whom correspondence should be addressed. Tel: +49 921 55 3540; Fax: +49 921 55 3544; Email: roesch@uni-bt.de

supported by the fact that NusE in a NusB deletion background could still support N antitermination and Nun termination (24). Although UV-crosslinking studies indicate that both NusB and NusE^{Δloop} contact *BoxA* RNA, detailed structural information about the RNA binding of NusB and the ternary NusB:NusE:RNA complex is not currently available.

Several *nutR BoxA* mutants abolish N-mediated antitermination or Nun-mediated transcription arrest *in vivo* and/or *in vitro*, namely *nutR BoxA5* (G35U) (26), *nutR BoxA16* (C38A) (27), *nutR BoxA* (U39G) (28). Oddly, the 9-bp transversion mutant *nutR BoxA69* has little effect on N-mediated antitermination except to make it NusB-independent, and it was proposed that NusB competed for *BoxA* binding with an as yet unidentified inhibitor of N activity (14).

NusB101 (D118N) presents an intriguing gain of function variant that suppresses a block in N-mediated antitermination by NusA1 (L183R) and NusE71 (A86D) at 42°C (29,30). NusB^{D118N} has enhanced affinity for *rrn* and λ *nut BoxA* (29); for example, NusB^{D118N}:NusE can be UV-crosslinked to *BoxA*-containing RNAs more efficiently than wt NusB:NusE (24). However, whether the increased affinity originates from a charge effect, from different direct contacts of the amino acid at position 118 to the RNA, or from a combination of effects is not clear.

In the present study, we used biophysical and genetic approaches to delineate identity elements of *nut* RNA that are recognized by NusB and by heterodimeric NusB:NusE complex. Furthermore, we studied NusB D118 mutants to clarify the role of this amino acid in NusB:RNA and NusB:NusE:RNA interactions.

MATERIALS AND METHODS

Cloning, expression and protein purification of Nus-factors

The *nusB* gene was cloned *via* BamH1 and Nde1 restriction sites into the *E. coli* expression vector pET29b (Novagen, Madison, WI, USA). *Escherichia coli* strain BL21(DE3) (Novagen, Madison, WI, USA) harboring the recombinant plasmid was grown at 37°C in LB (Luria-Bertani) medium containing kanamycin (30 µg/ml) until an OD₆₀₀ = 0.5 was reached, then the temperature was reduced to 20°C for 30 min and the cells were induced by 1 mM isopropyl 1-thio-β-D-galactopyranoside (IPTG). Cells were harvested 4 h after induction, resuspended in four times the pellet weight of lysis buffer (50 mM Tris, 150 mM NaCl, 1 mM DTT, pH 7.5), and lysed by using a micro-fluidizer (Microfluidics, Newton, MA, USA). After centrifugation the supernatant was dialyzed for 4 h against lysis buffer without NaCl and afterwards applied to a HeparinFF column (GE Healthcare, Munich, Germany) using a step gradient with increasing NaCl concentrations (0–1 M). For further purification the eluted fractions containing NusB were pooled and concentrated with Vivaspın concentrators (Vivascience, MWCO 5000 Da). The concentrated sample was applied to an S75 gel filtration column (GE Healthcare). The fractions containing NusB were pooled

and dialyzed against buffer as used for fluorescence measurements (25 mM HEPES, 100 mM potassium acetate, pH 7.5). The identity and structural integrity of the purified protein was analyzed by 19% SDS-PAGE and NMR spectroscopy.

NusB mutations

For NusB^{D118N} (NusB101), NusB^{D118R}, NusB^{D118A}, NusB^{D118E} and NusB^{D118K} the mutation primers shown in Supplementary Table S1 were used. Mutations were introduced by using the QuikChange protocol (Stratagene, La Jolla, CA, USA). Expression and purification was as described for wildtype NusB. NusB^{K2E} was inherent in the original NusB pETM11-plasmid (24).

NusE:NusB complex

NusE was cloned *via* BamH1 and EcoR1 restriction sites into the *E. coli* expression vector pGEX-6P (GE Healthcare) (24). The recombinant plasmid encoded a GST-NusE fusion protein with an internal PreScission cleavage site following the GST-tag. *Escherichia coli* strain BL21(DE3) (Novagen) harboring the recombinant plasmid was grown at 37°C in LB medium containing ampicillin (100 µg/ml) until an OD₆₀₀ = 0.5 was reached, then the temperature was reduced to 20°C for 30 min and the cells were induced by 1 mM IPTG. After induction overnight, the cells were harvested and resuspended in four times the pellet weight of lysis buffer (50 mM Tris, 150 mM NaCl, 1 mM DTT, pH 7.5). At this point, the NusB cell extract solved in the same buffer was added. After mixing for 20 min the cells were lysed with a micro-fluidizer (Microfluidics, Newton, MA, USA), and additional mixing was performed for 1 h to ensure correct formation of the NusB:NusE dimer. After centrifugation the dimer was purified from the supernatant *via* a GSTrap-FF column (GE Healthcare) using a one step elution (lysis buffer with 15 mM reduced glutathione). The GST-NusE fusion protein was cleaved by PreScission protease while dialyzing against lysis buffer at 4°C overnight. The cleaved protein was reapplied to a GSTrap-FF column using the same step elution procedure, but this time collecting the flow-through. For further purification the eluted fractions containing NusB:NusE were pooled and concentrated with Vivaspın concentrators (Vivascience, MWCO 5000 Da). The concentrated sample was applied to an S75 gel filtration column (GE Healthcare). The fractions containing NusB:NusE were pooled and dialyzed against buffer as used for fluorescence measurements (25 mM HEPES, 100 mM potassium acetate, pH 7.5). The identity and structural integrity of the purified protein complexes were analyzed by 19% SDS-PAGE.

NusB^{D118E}-NusE^{Δloop} production and purification for crystallization

Cloning of the genes encoding NusB and NusE^{Δloop} has been described (24). Mutations were introduced by using the QuikChange protocol (Stratagene, La Jolla, CA, USA). To produce protein for crystallographic analysis, plasmids containing the genes of interest were

co-transformed into *E. coli* strain BL21(DE3) by electroporation. The cells were grown in auto-inducing medium (31) in the presence of the appropriate antibiotics to an OD₆₀₀ of 0.5 at 37°C, and then incubated for an additional 16 h at 20°C. After harvesting at 4°C, the cell pellets were washed with binding buffer (50 mM Tris, pH 7.5, 150 mM NaCl) and stored at -80°C. Purification of the NusB^{D118N}:NusE^{Δloop} complex followed a double affinity chromatography procedure as described for the NusB:NusE^{Δloop} complex (24).

Crystallographic analysis

NusB^{D118N}:NusE^{Δloop} complex (NusB101:NusE^{Δloop}; 16 mg/ml) was crystallized at 20°C *via* the sitting drop vapor diffusion method by mixing 1 ml of sample with 1 ml of reservoir solution (0.2 M potassium citrate, 20% PEG 3350). Crystals could be flash frozen in liquid nitrogen after transfer into 60% reservoir plus 40% glycerol. Diffraction data were collected at 100 K on beamline PXII (SLS, Villigen, Switzerland) using a MarCCD 225 mm detector. The data were processed with the XDS package (32).

The structure of the NusB^{D118N}:NusE^{Δloop} complex was solved by molecular replacement using the coordinates of the NusB:NusE^{Δloop} complex [PDB ID 3D3B; (24)]. The model was manually rebuilt using COOT (33) and refined by standard methodology using Refmac5 including TLS refinement (34). Each protein molecule in the crystallographic asymmetric unit represented a separate TLS group.

In vivo assays

nusB::Cam nusA⁺ or *nusB::Cam nusA1* mutants carrying *λcI857* prophage were constructed. Wild-type and mutant NusB were supplied from a pBAD30 plasmid. Phage burst size after thermal induction was determined according to standard protocols.

Fluorescence equilibrium measurements

Various RNA sequences corresponding to the *nut* regions of the *λ* genome or to the *rrnG BoxA* of the *E. coli* genome (Supplementary Table S2) were used. Fluorescence equilibrium titrations were performed using an L-format Jobin-Yvon Horiba Fluoromax fluorimeter (Edison, NJ, USA). 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 × 4 mm quartz cuvette (Hellma, Müllheim, Germany). The excitation wavelength was 492 nm, and the emission intensity was measured at 516 nm applying a 500 nm cutoff filter. Anisotropic measurements were performed with slit widths of 4 nm and 3 nm for excitation and emission, respectively. All titration measurements were carried out at 25°C with 50 nM of 6-FAM-labeled RNA. Following sample equilibration, six data points with an integration time of 0.8 s were collected for each titration point.

Data fitting

Anisotropic data were fitted to a two-state binding model to determine the equilibrium dissociation constant (K_d) using standard software. The anisotropy was calculated from:

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

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

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

$$[\text{complex}] = \frac{(K_d + [P]_0 + [\text{RNA}]_0)}{(2[\text{RNA}]_0)} \quad 3$$

$$- \frac{\sqrt{(K_d + [P]_0 + [\text{RNA}]_0)^2 - 4[P]_0[\text{RNA}]_0}}{(2[\text{RNA}]_0)}$$

with A , anisotropy; A_{RNA} , initial free anisotropy; A_{complex} , anisotropy of the protein-RNA complex; P_0 , RNA₀, total protein and RNA concentration, respectively; R , ratio of intensities of bound and free forms.

CD measurements

Far UV CD measurements were performed on a J-810 S spectropolarimeter with a CDF-426S temperature control unit (JASCO International, Tokyo, Japan). Samples were prepared by dialyzing protein solutions against 10 mM sodium phosphate buffer, pH 7.5. Spectra were recorded at 25°C in a wavelength range of 185–260 nm with 50 nm/min scanning speed in a 1 mm path length quartz cuvette (Hellma, Müllheim, Germany) at a protein concentration of 10 μM. Buffer spectra were subtracted and ten spectra were accumulated. In order to normalize the measured ellipticity the mean residue molar ellipticity was calculated as:

$$[\Theta]_{\text{MRW}} = \frac{\Theta}{(c \cdot d \cdot N)} \quad 4$$

Θ , measured ellipticity; MRW, mean residue mass; c , protein concentration; d , path length; N , number of amino acids.

Thermal stability was analyzed by monitoring the CD signal at 222 nm during heating from 25°C to 90°C with a heating rate of 1°C/min. Quartz cuvettes with 1 cm path length equipped with a stirrer were used at a protein concentration of 2.5 μM. Both baselines and the transition region were fitted simultaneously:

$$y_{\text{obs}} = \frac{(y_n + m_n \cdot T_m)}{(1 + \exp(\Delta H_m/R(1/T_m - 1/T)))} + \frac{(y_d + m_d \cdot T) \cdot \exp(\Delta H_m/R \cdot (1/T_m - 1/T))}{(1 + \exp(\Delta H_m/R(1/T_m - 1/T)))} \quad 5$$

y_{obs} , observed ellipticity; y_n , y_d y-intercepts of the baselines of native and denatured protein; m_n , m_d , baseline slopes.

ΔH_m is the enthalpy at the temperature of the melting point (T_M) (35,36). Evaluations were based on the assumption that the unfolding transition is a two-state reaction. The temperature dependence of ΔG_D (free energy of unfolding) can be predicted at any temperature from the modified Gibbs–Helmholtz equation:

$$\Delta G_D(T) = \Delta H_m(1 - T/T_m) - \Delta C_p((T_m - T) + T \cdot \ln(T/T_m)) \quad 6$$

Over the narrow temperature range of the transition effects of ΔC_p (change in heat capacity) are negligible. Therefore, the equilibrium constant of the unfolding reaction, K , is defined as

$$-R \cdot T \ln(K) = \Delta H_m - T \cdot \Delta S_m = \Delta G_D \quad 7$$

$$K = \exp\left[\left(\frac{-\Delta H_m}{RT}\right) + \left(\frac{\Delta S_m}{R}\right)\right] \quad 8$$

where T is the temperature in Kelvin, R is the gas constant and ΔS_m is the entropy of unfolding at the melting point (T_M). When observing a two-state process with an experimental observable, y_{obs} , the equilibrium constant for the reaction is

$$K = \frac{(y_{\text{obs}} - (y_n + m_n \cdot T))}{((y_d + m_d \cdot T) - y_{\text{obs}})} \quad 9$$

By combining equations 7 and 9, ΔG_D can be calculated (36,37).

RESULTS

Fluorescence anisotropy measurements were performed to determine the dissociation constants of NusB or NusB:NusE and various RNA constructs (Tables 1 and 2).

Formation of the dimeric NusB:NusE complex enhances RNA binding affinity

NusB bound *rrnG BoxA-spacer* and *rrnG BoxA* with nearly identical efficiencies (dissociation constants K_d of 130 ± 20 nM and 200 ± 10 nM, respectively; Figure 1A, Table 1). The efficiency of NusB binding to λ *nutR*

BoxA-spacer, *nutL BoxA-spacer* and *nutR BoxA* sequences was significantly lower (K_d values of ~ 1.5 μ M; Figure 1B). The higher anisotropy of *nutR BoxA* relative to the other constructs reflects the different rotational correlation time relative to the large RNA constructs. Preformed NusB:NusE heterodimer bound to all RNAs tested with affinities more than an order of magnitude greater than NusB alone, consistent with the findings that both NusB and NusE in the NusB:NusE complex make RNA contacts (24). K_d values for the *rrnG BoxA-spacer* and the *rrnG BoxA* were 8 ± 2 and 5 ± 1 nM, respectively (Figure 1A; Table 2), and 90 ± 37 nM for *nutR BoxA-spacer*, 80 ± 22 nM for *nutL BoxA-spacer*, and 83 ± 8 nM for *nutR BoxA* (Figure 1C; Table 2). These data also indicate that contacts of NusE to the *spacer* region previously seen by UV-induced crosslinking (24) do not significantly increase the RNA affinity of the complex.

Protein RNA interaction takes place predominantly via BoxA

The binding of NusB and NusB:NusE to λ *nut BoxA-spacer* and *nutR BoxA* sequences with virtually identical affinities suggests that the spacer regions, shown previously to bind NusA (5), do not bind NusB or NusB:NusE. To confirm this, we measured binding to *spacer* alone. Specific binding of NusB or NusB:NusE to *nutL spacer* and *nutR spacer* was not observable in fluorescence titrations (Figure 1D). A slight increase of the fluorescence anisotropy signal with the NusB:NusE *nutR spacer* titration is consistent with unspecific binding with a K_d -value in the upper micromolar range.

BoxA mutations decrease binding affinity

Several *BoxA* mutations (Supplementary Table S2) affect λ N antitermination and HK022 Nun transcription arrest (14,26–28). We studied the effect of these mutations on NusB and NusB:NusE binding affinities. No interaction with the *boxA* transversion mutant *nutR BoxA69-spacer* and either NusB or NusB:E heterodimer was detected (Figure 2A and B; Tables 1 and 2), clearly indicating that the binding of these factors is *BoxA* RNA

Table 1. Dissociation constants for NusB monomer variants in nano molar

	NusB	NusB ^{D118N}	NusB ^{D118A}	NusB ^{D118E}	NusB ^{D118K}	NusB ^{D118R}
<i>rrnG BoxA-spacer</i>	130 ± 20	24 ± 4	230 ± 140	370 ± 30	240 ± 40	500 ± 130
<i>rrnG BoxA</i>	200 ± 10	50 ± 15	1100 ± 100	150 ± 20	450 ± 50	800 ± 80
<i>nutR BoxA-spacer</i>	1200 ± 600	900 ± 300	nb	3200 ± 800	nb	nb
<i>nutL BoxA-spacer</i>	2200 ± 800	1400 ± 500				
<i>nutR BoxA</i>	1600 ± 100	290 ± 15	500 ± 140	1400 ± 100	600 ± 25	1000 ± 70
<i>nutR-spacer</i>	nb	nb				
<i>nutL-spacer</i>	nb	nb				
<i>nutR BoxA5-spacer</i>	3100 ± 1600	12200 ± 1600				
<i>nutR BoxA16-spacer</i>	5100 ± 1400	6500 ± 2000				
<i>nutR BoxA(U39G)-spacer</i>	1600 ± 800	9000 ± 3400				
<i>nutR BoxA69-spacer</i>	nb	nb				

25 mM HEPES, 100 mM potassium acetate, pH 7.5.

nb = no binding detectable; empty cell = not determined. At least two independent experiments were performed per K_d . The relative molecular weights of the amino acids at residue 118 are A < N < D < E < K < R.

Table 2. Dissociation constants for NusB:NusE heterodimer variants in nano molar

	NusB:NusE	NusB ^{D118N} : NusE	NusB ^{D118A} : NusE	NusB ^{D118E} : NusE	NusB ^{D118K} : NusE	NusB ^{D118R} : NusE	NusB: NusE ^{Δloop}
<i>rrnG BoxA-spacer</i>	8 ± 2	24 ± 4	30 ± 5	125 ± 20	42 ± 6	26 ± 5	22 ± 2
<i>rrnG BoxA</i>	5 ± 1	25 ± 7	60 ± 5	135 ± 5	90 ± 4	65 ± 6	50 ± 5
<i>nutR BoxA-spacer</i>	90 ± 37	32 ± 11	700 ± 100	290 ± 90	110 ± 20	110 ± 28	360 ± 50
<i>nutL BoxA-spacer</i>	80 ± 22	40 ± 10					600 ± 60
<i>nutR BoxA</i>	83 ± 8	9 ± 1	60 ± 5	200 ± 20	75 ± 3	31 ± 2	160 ± 20
<i>nutR Spacer</i>	nb						
<i>nutL Spacer</i>	nb						
<i>nutR BoxA5-spacer</i>	5200 ± 1300						
<i>nutR BoxA16-spacer</i>	3400 ± 1800						
<i>nutR BoxA(U39G)-spacer</i>	1900 ± 700						
<i>nutR BoxA69-spacer</i>	nb						

25 mM HEPES, 100 mM potassium acetate, pH 7.5.

nb = no binding detectable; empty cell = not determined. At least two independent experiments were performed per K_d . The relative molecular weights of the amino acids at residue 118 are A < N < D < E < K < R.

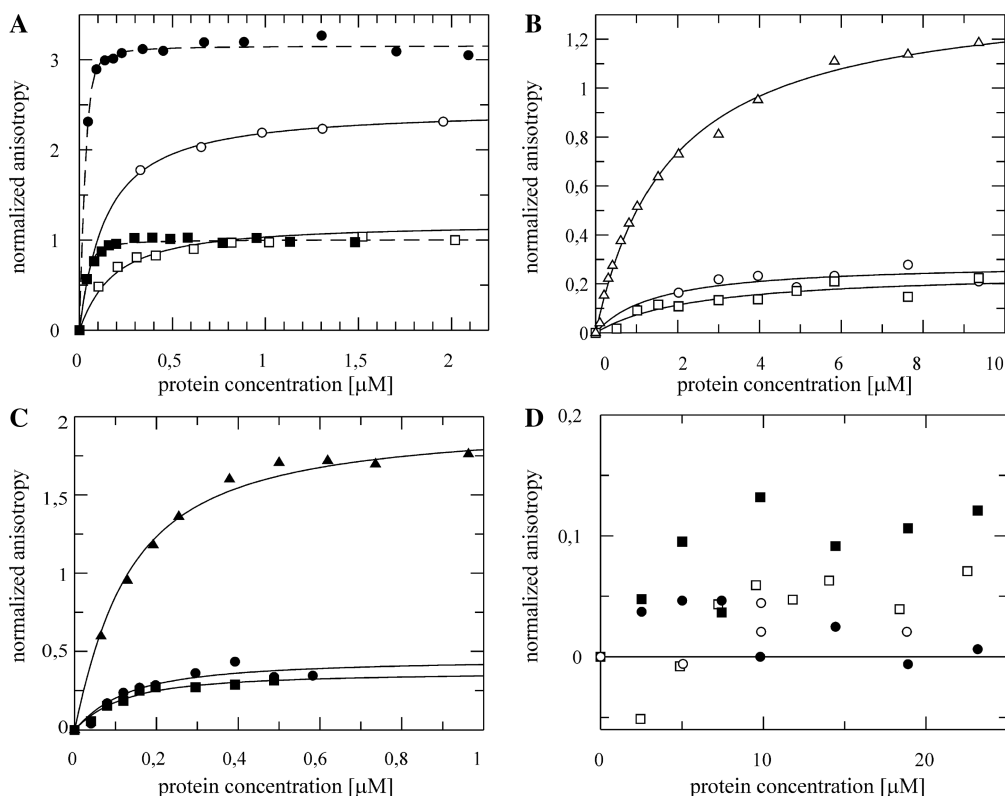


Figure 1. Fluorescence anisotropy titration of fluorescein labeled λ *nut* and *rrnG* RNAs with NusB (open markers) and NusB:NusE complex (filled markers). (A) A 50 nM *rrnG BoxA-spacer* (squares) and 50 nM *rrnG BoxA* (circles) titrated with NusB and NusB:NusE. (B) Fifty nanomolar *nutR BoxA-spacer* (circles), *nutL BoxA-spacer* (squares) and *nutR BoxA* (triangles) titrated with NusB. (C) A 50 nM *nutR BoxA-spacer* (circles), *nutL BoxA-spacer* (squares) and *nutR BoxA* (triangles) titrated with NusB/NusE. (D) A 50 nM *nutR spacer* (circles) and 50 nM *nutL spacer* (squares) titrated with NusB and NusB/NusE. Solid lines represent the best fit to equation (3).

sequence-dependent. This result supports the *in vivo* finding of Patterson *et al.* (14) that N antitermination on *BoxA69* fusions is NusB-independent. We next tested *BoxA* point mutants that inhibit N and Nun (Figure 2A and B; Tables 1 and 2). The *nutR BoxA5* mutation (G35U) reduced NusB binding 2- to 3-fold ($K_d = 3.1 \pm 1.6 \mu\text{M}$). Interestingly, the affinity of NusB:NusE heterodimer for the mutant RNA was essentially identical to that of NusB

($K_d = 5.2 \pm 1.3 \mu\text{M}$). NusB and NusB:NusE bound *nutR BoxA16* (C38A) with K_d values of $5.1 \pm 1.4 \mu\text{M}$ and $3.4 \pm 1.8 \mu\text{M}$, respectively. The dissociation constants of NusB and NusB:NusE for *nutR BoxA* (U39G), which inhibits N, were $1.6 \pm 0.8 \mu\text{M}$ and $1.9 \pm 0.7 \mu\text{M}$, respectively. Judged from their effects on the NusB and NusB:NusE binding affinities, G35 and C38 participate more tightly in protein binding than U39. The equivalent

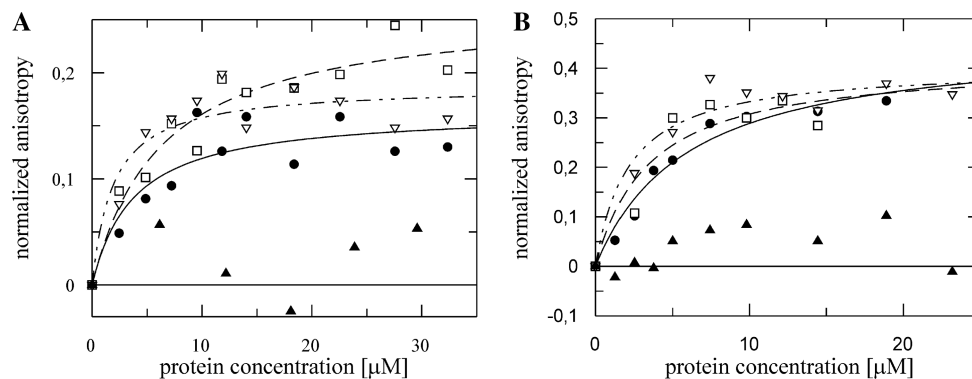


Figure 2. Fluorescence anisotropy, λ *nut BoxA* variants and NusB (A), NusB:NusE (B). Each titration was performed with 50 nM of fluorescein labeled RNA. *nutR BoxA5-spacer* (filled circle, fit: solid line), *nutR BoxA16-spacer* (open square, fit: dashed line), *nutR BoxA69-spacer* (filled triangle) and *nutR BoxA(U39G)-spacer* (open triangle, fit: dotted line). Lines represent the best fit to equation (3). For *nutR BoxA69-spacer* no interaction was observable, therefore no fitting was performed.

binding affinities of NusB and NusB:NusE heterodimer to these *BoxA* point mutants stands in sharp contrast to the enhanced binding of the heterodimer to wild-type *BoxA* sequences.

The NusB^{D118N} (NusB101) mutation affects RNA binding affinity

NusB^{D118N} is a gain-of-function mutation that enables NusB to override mutations in NusE and NusA that abrogate λ N-mediated antitermination (29,30). The dissociation constants for NusB^{D118N} complexes with *rrnG BoxA-spacer* and *rrnG BoxA* were 5- to 6-fold lower compared to wild-type NusB (Figure 3A; Table 3). Only small reductions in K_d were observed for the *nutL BoxA-spacer* and *nutR BoxA-spacer* complexes, whereas the dissociation constant for the complex of NusB^{D118N} with *nutR BoxA* decreased significantly from ~1600 nM to ~290 nM (Figure 3B; Table 1). In contrast, the mutation increased the K_d values about 3-fold for heterodimer complexes with *rrnG BoxA-spacer* and *rrnG BoxA* (Figure 3C, Table 2). NusB^{D118N}:NusE complexes with *nutR BoxA-spacer* and *nutL BoxA-spacer* sequences displayed 3- and 5-fold, respectively, lower K_d values for the mutant relative to the wild-type heterodimer (Figure 3D, Table 3). For the NusB^{D118N}:NusE complex with *nutR BoxA*, the K_d decreased by nearly an order of magnitude to 9 ± 1 nM (Figure 3D; Table 2). Thus, NusE did not further enhance the binding of NusB^{D118N} to *rrnG BoxA* or *rrnG BoxA-spacer*. However, NusE strongly stimulated the binding of NusB^{D118N} to sequences derived from *nutL* and *nutR*.

The structure of NusB^{D118N}:NusE ^{Δ loop} closely resembles the structure of NusB:NusE ^{Δ loop}

NusE ^{Δ loop} is a derivative of NusE that binds NusB and retains transcriptional but not translational activity (24). The NusB:NusE ^{Δ loop} complex binds RNA that includes a *BoxA* sequence, although with lower efficiency than NusB:NusE [(24); Figure S1; Table 2]. Previous studies demonstrated that NusB^{D118N}:NusE ^{Δ loop} bound RNA more tightly than NusB:NusE ^{Δ loop}. To ask if the increased

RNA affinity of the NusB^{D118N}:NusE ^{Δ loop} complex correlated with structural rearrangements compared to the NusB:NusE ^{Δ loop} complex, we solved the crystal structure of the NusB^{D118N}:NusE ^{Δ loop} complex by molecular replacement at 2.5 Å resolution (Figure 4). The structure was refined to R_{work} and R_{free} factors of 20.4% and 25.6%, respectively (Table 4). An asymmetric unit of the crystal contained three molecules each of NusB^{D118N} and NusE ^{Δ loop}, which formed three NusB^{D118N}:NusE ^{Δ loop} complexes. Two of these complexes exhibited well-defined electron density, but the electron density map of the third complex was fragmentary: In that complex, residues 60–77 and 127–139 of NusB^{D118N} and residues 45–47 and 60–72 of NusE ^{Δ loop} could not be unambiguously traced. The following discussion therefore refers to the structures of the two well defined complexes, which closely resemble each other [RMSD of 0.75 Å for 220 C α atoms; calculated with SSM (38)].

The global structure of NusB^{D118N} in complex with NusE ^{Δ loop} is very similar to that of wild-type NusB in isolation [PDB ID 1EY1; (39); rmsd of 2.54 Å for 110 C α atoms; Figure 4]. Furthermore, the structure of the NusB^{D118N}:NusE ^{Δ loop} complex is virtually identical to that of the NusB:NusE ^{Δ loop} complex (RMSD of 0.85 Å for 220 C α atoms; Figure 4B), demonstrating that the D118N mutation has no global conformational consequences. In particular, the positions and conformations of NusB residue N118 in the mutant and of residue D118 in the parent complex are essentially identical. Irrespective of the amino acid at position 118, the neighboring region undergoes identical adjustments upon NusE ^{Δ loop} binding, during which the C α position of residue 118 is repositioned by 2.8 Å (Figure 4C, inset). However, the D118N exchange induces a significant difference in the local electrostatic surface properties of the complex (Figure 4D). This observation is consistent with the idea that the increased RNA affinity of NusB^{D118N} or its complex with NusE ^{Δ loop} is at least in part due to the replacement of a negatively charged residue with an uncharged residue at the RNA binding site, thus reducing repulsion with the negatively charged

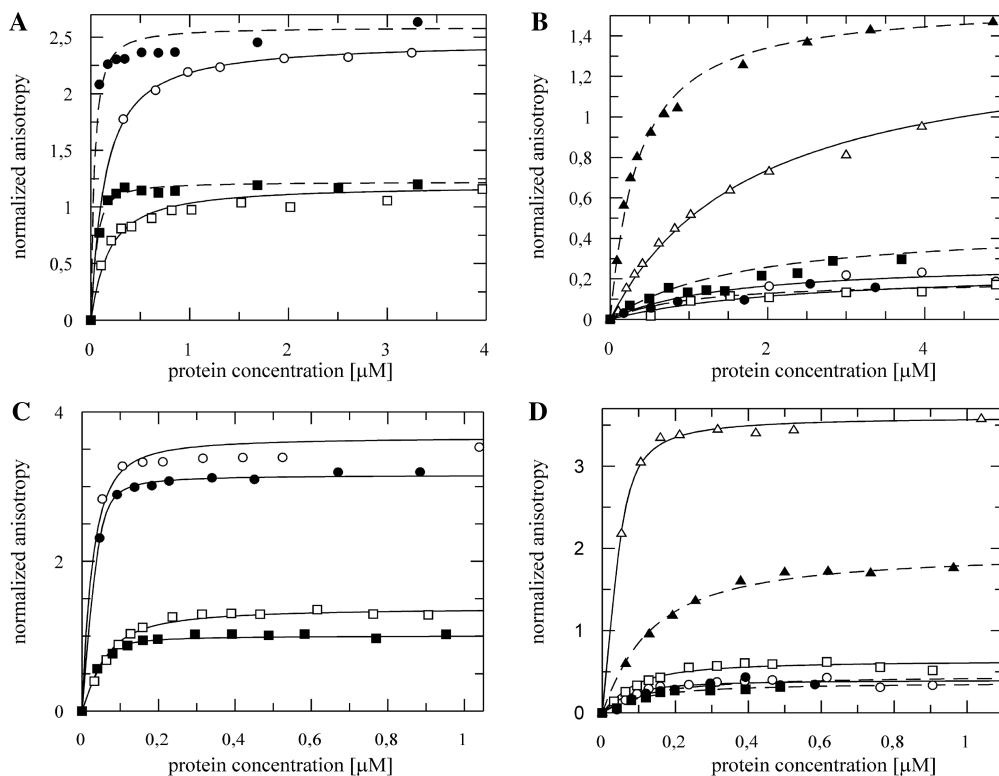


Figure 3. Fluorescence anisotropy measurements with fluorescein-labeled *λ**nut* and *rrnG* RNAs towards NusB^{D118N} (A and B) and NusB^{D118N}:NusE complex (C and D) (filled markers) compared to the wild-type NusB and NusB:NusE complex (open markers, data as in Figure 1A and C). (A) Fifty nanomolar *rrnG* *BoxA-spacer* (squares) and 50 nM *rrnG* *BoxA* (circles) titrated with NusB^{D118N} and NusB; (B) 50 nM *nutR* *BoxA-spacer* (circles), *nutL* *BoxA-spacer* (squares) and *nutR* *BoxA* alone (triangles) titrated with Nus^{D118N} and NusB. (C) A 50 nM *rrnG* *BoxA-spacer* (squares) and 50 nM *rrnG* *BoxA* (circles) titrated NusB^{D118N}:NusE and NusB:NusE; (D) 50 nM *nutR* *BoxA-spacer* (circles), *nutL* *BoxA-spacer* (squares) and *nutR* *BoxA* (triangles) titrated with NusB^{D118N}:NusE and NusB:NusE. Dashed lines represent the best fit to equation (3) for NusB^{D118N} and NusB^{D118N}/NusE, respectively. Solid lines the similar fit for wild-type NusB and NusB:NusE.

Table 3. Melting temperatures (T_M), free reaction enthalpy at the melting point ($\Delta H_{M,D}$) and Gibbs free energy of the unfolding reaction at 328K (ΔG_D) values for NusB variants (10 mM potassium phosphate, pH 7.5)

	T_M (K)	$\Delta H_{M,D}[T_M]$ (kJ/mol)	ΔG_D [328K] (kJ/mol)
NusB	337.8 ± 0.1	280 ± 4	8.3 ± 0.2
NusB ^{D118N}	333.6 ± 0.1	174 ± 3	2.7 ± 0.2
NusB ^{D118A}	333.5 ± 0.1	340 ± 4	5.6 ± 0.2
NusB ^{D118R}	331.0 ± 0.1	298 ± 4	2.4 ± 0.2
NusB ^{D118E}	334.5 ± 0.1	263 ± 3	5.8 ± 0.2
NusB ^{D118K}	330.8 ± 0.1	244 ± 3	1.6 ± 0.2

sugar-phosphate backbone of the RNA. Alternatively, or in addition, introduction of an asparagine for an aspartate at position 118 may result in additional hydrogen bonds to the RNA.

The overall structures of various NusB118 mutants are highly similar

To investigate further the effect of amino acid variations at position 118, several point mutations with positively charged, negatively charged, and apolar amino acids were examined. All NusB variants show CD-spectra typical of α -helices, i.e. minima at 208 nm and 222 nm as

well as a maximum at 190 nm (Figure 5A), with only minor differences from the wild-type protein spectrum. The melting temperatures of NusB and NusB D118 variants were determined by thermal unfolding. The CD signal at 222 nm, which we used to indicate melting, was reduced by all mutations, particularly by substitutions with positively charged residues. The Gibbs free energy of the unfolded species (ΔG_D) at 328 K thus ranges from 8.3 ± 0.2 kJ/mol for wild-type NusB to 1.6 ± 0.2 kJ/mol for NusB^{D118K}. The NusB variants, with the notable exception of NusB^{D118N}, show very similar unfolding transitions. The broader transition of NusB^{D118N} indicates a lower value for the Gibbs free energy of the unfolding reaction at 328 K and for the free reaction enthalpy at the melting point [(35,36); Figure 5 and Table 3]. These relatively small differences indicate that the D118 point mutations do not lead to global NusB misfolding or to unstable NusB proteins.

Effects of other NusB D118 mutations on RNA binding

The binding properties of different NusB mutants dependent on the RNA were tested. Thus the binding of NusB^{D118A} to *rrnG* *BoxA-spacer* was approximately as tight as NusB⁺, whereas the affinity of the mutant for *rrnG* *BoxA* was 20% that of NusB⁺ (Figure 6; Table 1). NusB^{D118E} bound *rrnG* *BoxA* with wild-type efficiency but

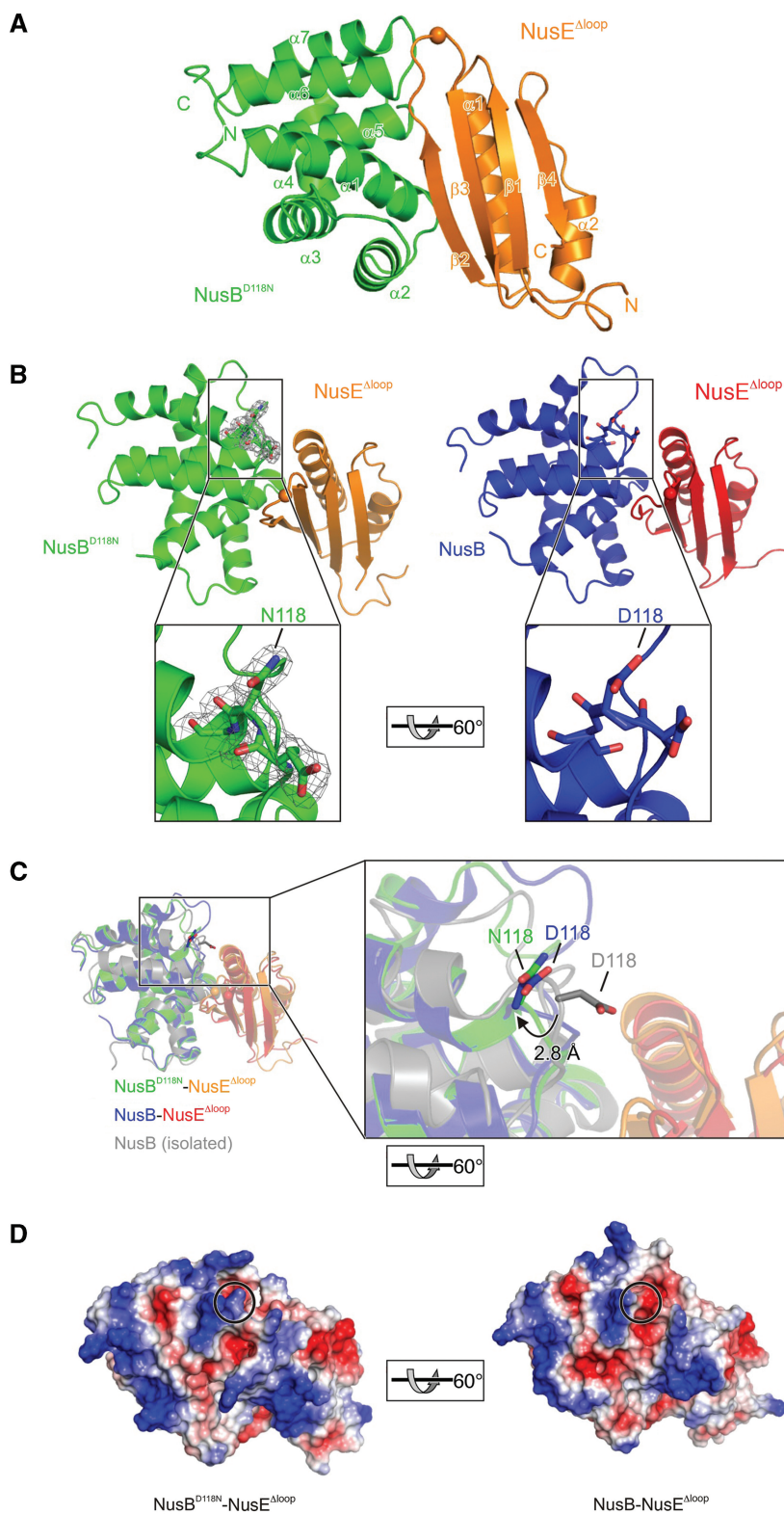


Figure 4. Structure of the NusB^{D118N}::NusE^{Δloop} complex. (A) Ribbon plot of the *E. coli* NusB^{D118N}::NusE^{Δloop} complex. NusB^{D118N}, green; NusE^{Δloop}, orange. Secondary structure elements and termini are labeled. The orange sphere marks the site at which the ribosome-binding loop of NusE has been replaced by a single serine. (B) Comparison of the NusB^{D118N}::NusE^{Δloop} complex (left) with the NusB::NusE^{Δloop} complex [right, PDB ID 3D3B; (24)]. Insets: closeup views of the residue 118 regions. The orientation relative to (A) is indicated. Gray mesh, final $2F_o - F_c$ electron density of the NusB^{D118N}::NusE^{Δloop} structure contoured at the 1σ level and covering N118 and neighboring residues. The orientation relative to (A) is indicated. (C) Superimposition of the NusB::NusE^{Δloop} complex [blue and red, PDB ID 3D3B; (24)] and of NusB [grey, PDB ID 1EY1; (39)] on the NusB^{D118N}::NusE^{Δloop} complex (green and orange). Residues at position 118 are shown as sticks and a magnified view of the residue 118 region is provided (carbon, as the respective molecule; oxygen, red; nitrogen, blue). The orientation relative to (A) is indicated. (D) Comparison of the electrostatic surface potentials of the complexes. Blue, positive charge; red, negative charge. Left, NusB^{D118N}::NusE^{Δloop} complex. Right, NusB::NusE^{Δloop} complex. The positions of residue 118 are circled. The orientations are the same as in (B).

Table 4. Crystallographic data

	NusB ^{D118N} – NusE ^{Δloop}
Data collection	
Wavelength (Å)	0.9788
Temperature (K)	100
Space group	I4 ₁ 22
Unit cell parameters (Å, °)	<i>a</i> = 112.64, <i>b</i> = 112.64, <i>c</i> = 263.25
Resolution (Å)	30.0–2.5 (2.6–2.5) ^a
Reflections	
Unique	29 761 (3263)
Completeness (%)	100 (100)
Redundancy	7.22 (7.42)
<i>I</i> / σ (<i>I</i>)	18.1 (4.1)
<i>R</i> _{sym} (<i>I</i>) ^b	8.6 (72.5)
Refinement	
Resolution (Å)	30.0–2.5 (2.56–2.50)
Reflections	
Number	29 753 (2178)
Completeness (%)	100 (100)
Test set (%)	5.0
<i>R</i> _{work} ^c	20.4 (23.1)
<i>R</i> _{free} ^c	25.6 (29.1)
Contents of AU ^d	
Protein molecules/refined atoms	3 NusB ^{Asp118Asn} , 3 NusE ^{Δloop} /5313
Water oxygen	155
Ions	1 K ⁺
Mean B-factors (Å ²)	
Wilson	52.7
Protein	60.5
Water	22.8
Ions	29.5
Ramachandran plot ^e	
Favored (%)	97.12
Allowed (%)	2.43
Outliers (%)	0.45
RMSD from target geometry	
Bond lengths (Å)	0.01
Bond angles (°)	1.22
RMSD B-factors (Å ²)	
Main chain bonds	0.41
Main chain angles	0.82
Side chain bonds	1.43
Side chain angles	2.48
PDB ID	3IMQ

^aData for the highest resolution shell in parentheses

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

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

^dAU, asymmetric unit.

^eCalculated with MolProbity (<http://molprobity.biochem.duke.edu/>) (40) RMSD, root-mean-square deviation.

associated with *rrnG BoxA-spacer* ~2-fold less well than NusB⁺. We considered the possibility that replacing D118 with a positively charged residue might enhance binding through ionic interactions with the RNA ligand. This was not the case. Neither NusB^{D118K} nor NusB^{D118R} bound *rrnG* RNA with wild-type efficiency. The higher volumes of the lysine and arginine residues may induce unfavorable steric interactions, canceling out positive contributions to binding by electrostatic interactions with RNA. Consistent with this idea is our observation that NusB^{D118R} bound *rrn* or *nutR BoxA* less efficiently than

NusB^{D118K}, reflecting, perhaps, the larger volume of the former substitution. Of the various D118 substitutions, interaction with the *nutR BoxA-spacer* could be detected only for NusB^{D118N} and NusB^{D118E} (Figure 6 and Table 3). NusB^{D118E} bound *nutR BoxA* as well as NusB⁺. NusB^{D118A} and NusB^{D118K} bound *nutR BoxA* significantly better than wild-type NusB (Figure 6 and Tables 1 and 2).

NusE improves binding of NusB mutants

With the exception of NusB^{D118E}, NusE significantly enhanced mutant NusB binding to *rrnG* RNA, although no mutant except NusB^{D118N} bound as well as NusB⁺: NusE. NusE also enhanced binding to *nutR* sequences. Binding of NusB mutants to *nutR BoxA* was, except for NusB^{D118E}, at least as strong as NusB⁺. Thus, NusE may foster additional RNA contacts, rendering NusB-mediated interactions less dominant (Figure 7 and Tables 1 and 2).

NusB^{D118N}, NusB^{D118K} and NusB^{D118R} suppress the *nusA1* (*nusA*^{L183R}) mutation *in vivo*

We next tested the NusB D118 mutants for suppression of *nusA1* (*nusA*^{L183R}). *nusA*^{L183R} prevents phage λ growth at 42°C by blocking λ N antitermination. Over-expression of NusB D118 mutants in a *nusA*⁺ λ *cI857* lysogen had modest negative effects on phage burst size (Table 5). The negatively charged NusB^{D118E} was most inhibitory, reducing burst size to 32% of wild-type levels. In the *nusA*^{L183R} background, all mutants except NusB^{D118E} increased burst size >100-fold. NusB^{D118N} suppressed *nusA*^{L183R} with greatest efficiency, increasing burst size from <0.01% to 4.9% relative to *nusA*⁺. NusB^{D118K} and NusB^{D118R} also significantly enhanced λ growth (to 2.2% and 3.3%, respectively, of *nusA*⁺ titers). Suppression by NusB^{D118A} was the least effective (0.5%). Taken together, these data suggest that replacing the negatively charged asp118 with an uncharged asparagine residue or a positively charged lysine or arginine residue significantly restores λ N activity in a *nusA*^{L183R} strain. Poor suppression by the alanine substitution may reflect the lower molecular weight of this aminoacid relative to aspartate. Similarly, the inability of NusB^{D118E} to restore λ growth may also be caused by steric effects due to the bulky glutamate side chain, in spite of its negative charge.

DISCUSSION

Processive transcription antitermination depends on formation of a multi-factorial ribonucleoprotein complex on the surface of RNAP in response to *nut* signaling sequences in the untranslated leader regions of transcripts. These factors include NusA, NusB, NusE and NusG (2–4). In the case of λ *nut*, λ N protein forms part of the complex, whereas the complex that forms at *rrn nut* includes the Nus proteins and a number of additional host factors (21,22). Protein–protein and protein–RNA interactions in these complexes are cooperative in the sense that a stable complex is assembled based on numerous, but often weak, binary interactions. While some

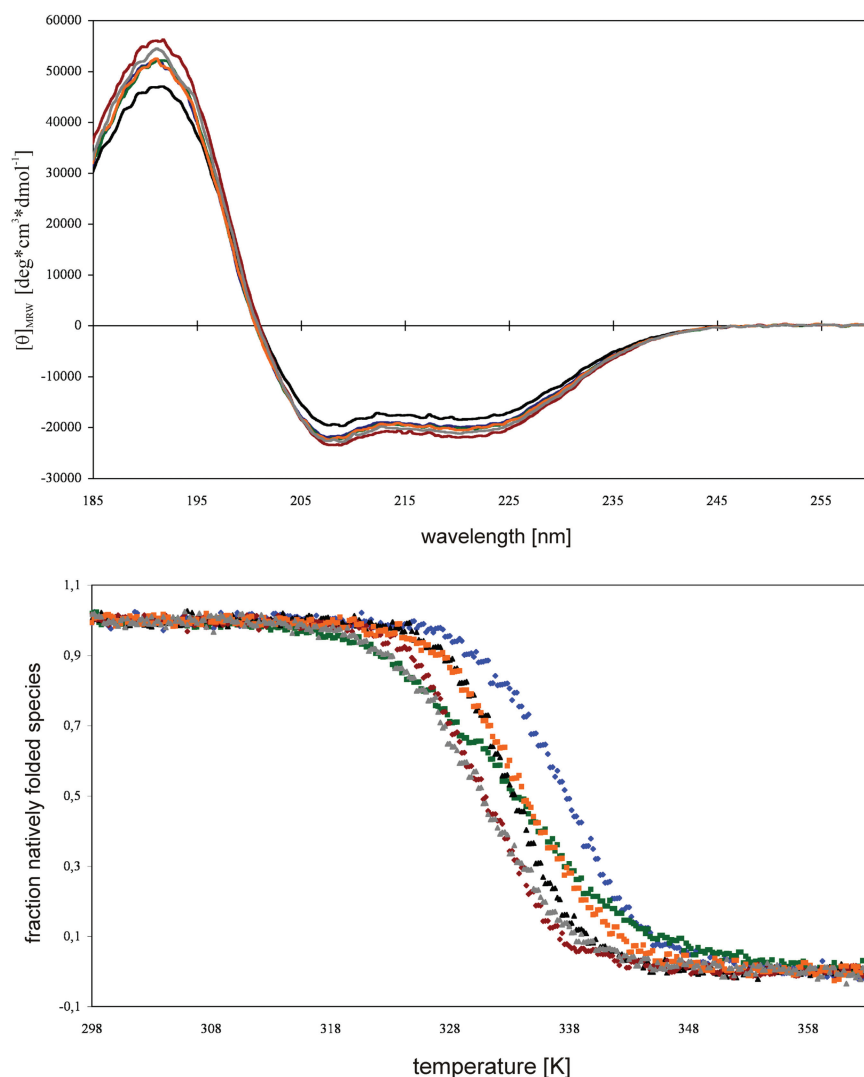


Figure 5. (A) Overlay of the CD-spectra for the NusB variants. Θ_{MRW} versus wavelength in nm. Θ_{MRW} was according to equation (4). (B) Thermal unfolding for the NusB variants. Fraction of the unfolded species versus absolute temperature. NusB (blue), NusB^{D118N} (green), NusB^{D118A} (black), NusB^{D118R}, (red), NusB^{D118E} (orange) and NusB^{D118K} (grey).

of these binary interactions have been mapped within the complexes, little is known about their relative strengths and how important these are for generating functional complexes. Here, we have used a combination of biophysical and functional studies to explore interactions between NusB or NusB:NusE complex and the *nut BoxA* RNA signaling sequence to which it binds.

We confirm that the affinity of NusB:NusE for *BoxA* is an order of magnitude higher than that of NusB alone (13), reflecting, presumably, the additional contacts that NusE makes with *BoxA*. We find that mutations in λ *nut BoxA* that inhibit λ N antitermination reduce NusB binding. However, in contrast to λ *nut BoxA*⁺, the affinities of NusB and NusB:NusE for the mutant *BoxA* sequences are essentially identical. The mutations lie throughout *BoxA* and do not define the known sites of NusE–RNA interactions (24).

The K_d values reported above for the association of NusB or NusB:NusE with λ *nut BoxA* are similar

to those reported by Greive and coworkers based on fluorescence anisotropy experiments (13). We note, however, a large difference with respect to *rrnG BoxA-spacer* binding. These authors report a K_d of 850 nM for NusB and 200 nM for NusB:NusE, whereas we observe values of 130 nM and 8 nM, respectively. Since the RNA sequences tested were identical, we suggest that the lower affinity seen by Greive *et al.* reflects the 5' location of the fluorescent label on their RNA compared to the 3' location used in our experiments. The N-terminus of NusB contacts both *rrn BoxA* and *nut BoxA* at their 5'-ends, and interference with these contacts by a 5' label could increase the K_d values, although in an RNA sequence-dependent manner. Indeed, we find that the K2E mutation increased the K_d of the NusB *rrn BoxA-spacer* complex from 130 nM to 3600 nM, but raised the K_d of NusB binding to *nutL BoxA-spacer* only 2-fold (from 2200 nM to 5100 nM).

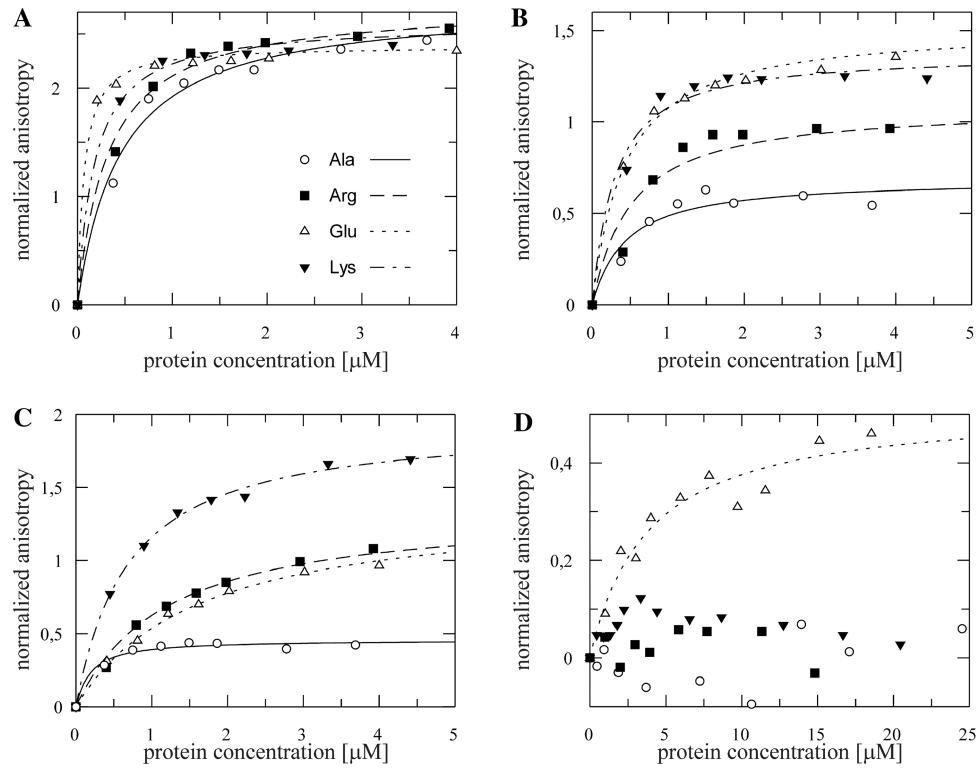


Figure 6. Fluorescence anisotropy measurements with fluorescein labeled *λnut* or *rrnG* RNAs and NusB^{D118A} (open circle, fit: bold line), NusB^{D118R} (filled square, fit: dashed line), NusB^{D118E} (open triangle, fit: pointed line) and NusB^{D118K} (filled triangle, fit: dashed/pointed line). Lines represent the best fit to equation (3). (A) Fifty nanomolar *rrnG* *BoxA*; (B) 50 nM *rrnG* *BoxA-spacer*; (C) 50 nM *nutR* *BoxA*; and (D) 50 nM *nutR* *BoxA-spacer*.

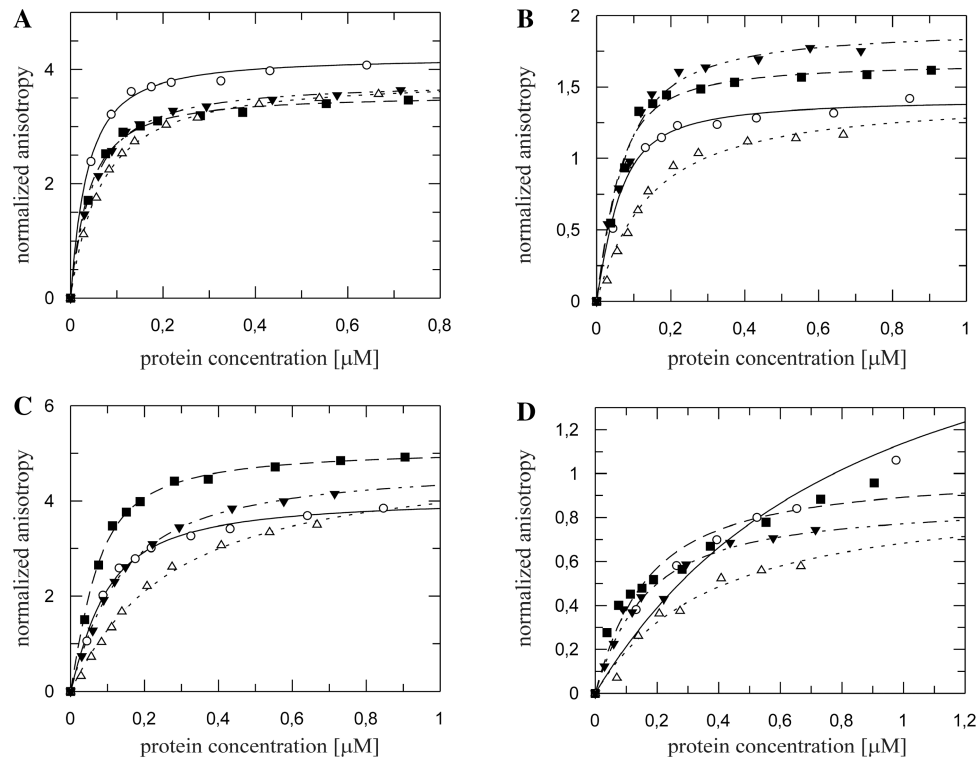


Figure 7. Fluorescence anisotropy measurements with fluorescein labeled *λnut* and *rrnG* RNAs to NusB^{D118A}:NusE (open circle, fit: solid line), NusB^{D118R}:NusE (filled square, fit: dashed line), NusB^{D118E}:NusE (open triangle, fit: pointed line) and NusB^{D118K}:NusE (filled triangle, fit: dashed/pointed line). Lines represent the best fit to equation (3). (A) Fifty nanomolar *rrnG* *BoxA*; (B) 50 nM *rrnG* *BoxA-spacer*; (C) 50 nM *nutR* *BoxA*; and (D) 50 nM *nutR* *BoxA-spacer*.

Table 5. Strains are W3110 *nusB::Cam λcI857* lysogens

<i>nusB</i> plasmid	<i>nusA</i> ⁺	<i>nusAI</i>	Suppression (%)
–	103	0.0004	<0.01
+	137	2	<0.01
D118N	85	4.2	4.9
D118A	74	0.4	0.5
D118E	44	0.005	<0.01
D118K	113	2.5	2.2
D118R	125	4.2	3.3

NusB is carried on pBAD30 and induced with 0.2% arabinose. Cells were grown overnight at 32°C in LB + ampicillin (50 µg/ml), diluted 1:100 into the same medium + 0.2% arabinose, and grown at 32°C for 1 h. λ was induced by temperature shift to 42°C for 90 min. Burst size was determined by plating the lysate on W3110 at 37°C. Cells were titered at 32°C prior to temperature shift; titers were equivalent for all strains. Values represent an average of two experiments; variation was <8 %.

We have also explored the phenotype of the NusB^{D118N} (NusB101) mutation. NusB^{D118N} suppresses mutations in NusA and NusE that inhibit λ N antitermination function (29,30). It has been proposed that an increase in the affinity of NusB^{D118N} for *nut* RNA compared to NusB⁺ may compensate for weaker RNA (or protein) contacts of the mutant Nus proteins (24, 29). We show here that the D118N mutation enhances NusB binding to both *rrn BoxA* and λ *nut BoxA*. The mutation reduces binding to λ *nut BoxA* mutants, *BoxA5* and *BoxA(U39G)*, but has no effect on binding to *BoxA16*. It will be interesting to test whether the D118N mutation further reduces λ N antitermination on *BoxA* mutant templates *in vivo*. D118N also enhances NusB binding to *nut BoxA* in complex with NusE. However, we find that the affinity of NusB^{D118N}:NusE for *rrn BoxA* is only 20% that of NusB⁺:NusE. These data imply that D118N, although it optimizes the formation of the λ N antitermination complex, interferes with the assembly of antitermination complexes on *rrn* operons.

The enhanced affinity of NusB^{D118N} to *BoxA* is not due to gross distortions in the shape of the NusE complex. Thus, the structure of NusB^{D118N}:NusE^{Δloop} complex is virtually identical to NusB⁺:NusE^{Δloop}, although the electrostatic surface properties are strongly affected (Figure 4D). In fact, CD spectroscopic analyses (Figure 5) show that NusB also structurally tolerates a number of other mutations at position 118. Thus, the effects of the D118 mutations on RNA binding are local.

D118 is located close to NusB residues that interact with *BoxA* RNA. It was suggested that removal of the negatively charged aspartate with the neutral asparagine residue might extend the NusB RNA-binding surface and that this might stabilize the antitermination complex and account for the suppression of *nusAI* and *nusE71*. Our results in general support this notion, although we find that the size of the D118 substitution also plays a role in the affinity of NusB for RNA.

The differential RNA affinities of the mutant NusB:NusE complexes roughly correlate with their *in vivo* suppression activity. We measured the burst sizes after induction of a λ *cI857* lysogen in *nusA*⁺ and *nusAI*

hosts. This assay, which reflects the ability of λ N to antiterminate, showed that the order of *nusAI* suppression efficiency was D118N > D118R > D118K > D118A. D118E failed to increase λ burst size over that seen for NusB⁺. The binding affinities to *nut BoxA* for the mutants in complex with NusE was D118N > D118R > D118A > D118K > D118E. Note that the *K_d* of NusB⁺ for *nut BoxA* (83 nM) was not significantly different from that of NusB^{D118K} (75 nM), yet the mutant NusB increased the λ burst size in a *nusAI* host at least 100-fold above the wild-type NusB level. This disjunction between RNA binding and N activation is, as yet, unexplained. It implies, however, that D118 may make functionally important contacts within NusB or with NusE.

We propose that D118 does not contact *BoxA*, but that removal of the negative charge permits such interaction. The location of position 118 opposite the NusB:NusE interaction surface renders an effect of mutations at this position on the NusB:NusE interaction unlikely. However, it cannot be ruled out completely, and a more definitive verdict needs further structural analysis of the NusB:NusE:*BoxA* complex. Study of how NusB mutations affect complex formation with NusE^{Δloop} and alter its RNA binding properties might shed further light on the role of the NusE loop in RNA binding.

The results from our combined biophysical and genetic investigations illustrate how a particular protein–RNA interaction is fine-tuned with respect to other interactions within a functional ribonucleoprotein complex to achieve processive transcriptional antitermination. Furthermore, our results refine the mechanism by which NusB acts as a NusE RNA loading factor (24). NusB:NusE:RNA complex formation is entirely mediated by the *BoxA* sequences, whereas the *spacer* regions are necessary and sufficient for NusA binding (5). Sterically, simultaneous binding of these Nus factors to *nut* and *rrn* should be possible. There is, however, no evidence that NusA and NusB:NusE interact at *nut*, and no increase in *nut* binding by NusB:NusE was observed on NusA addition (15).

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We would like to thank Ramona Heissmann for excellent technical assistance.

FUNDING

Deutsche Forschungsgemeinschaft (Ro617/16-1 to P.R.; Wa1126/3-1 to M.C.W.); National Institutes of Health (GM037219 to M.E.G.). Funding for open access charge: Universität Bayreuth.

Conflict of interest statement. None declared.

REFERENCES

- Greenblatt, J., Nodwell, J.R. and Mason, S.W. (1993) Transcriptional antitermination. *Nature*, **364**, 401–406.
- Das, A. (1993) Control of transcription termination by RNA-binding proteins. *Annu. Rev. Biochem.*, **62**, 893–930.
- Borukhov, S., Lee, J. and Laptenko, O. (2005) Bacterial transcription elongation factors: New insights into molecular mechanism of action. *Mol. Microbiol.*, **55**, 1315–1324.
- Roberts, J.W., Shankar, S. and Filter, J.J. (2008) RNA polymerase elongation factors. *Annu. Rev. Microbiol.*, **62**, 211–233.
- Prasch, S., Jurk, M., Washburn, R.S., Gottesman, M.E., Wöhrl, B.M. and Rösch, P. (2009) RNA-binding specificity of *E. coli* NusA. *Nucleic Acids Res.*, **37**, 4736–4742.
- Chattopadhyay, S., Garcia-Mena, J., DeVito, J., Wolska, K. and Das, A. (1995) Bipartite function of a small RNA hairpin in transcription antitermination in bacteriophage λ . *Proc. Natl Acad. Sci. USA*, **92**, 4061–4065.
- Robert, J., Sloan, S.B., Weisberg, R.A., Gottesman, M.E., Robledo, R. and Harbrecht, D. (1987) The remarkable specificity of a new transcription termination factor suggests that the mechanisms of termination and antitermination are similar. *Cell*, **51**, 483–492.
- Prasch, S., Schwarz, S., Eisenmann, A., Wöhrl, B.M., Schweimer, K. and Rösch, P. (2006) Interaction of the intrinsically unstructured phage λ N protein with *E. coli* NusA. *Biochemistry*, **45**, 4542–4549.
- Bonin, I., Mühlberger, 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 λ . *Proc. Natl Acad. Sci. USA*, **101**, 13762–13767.
- Nodwell, J.R. and Greenblatt, J. (1993) Recognition of *boxA* antiterminator RNA by the *E. coli* antitermination factors NusB and ribosomal protein S10. *Cell*, **72**, 261–268.
- Lüttgen, H., Robelek, R., Mühlberger, R., Diercks, T., Schuster, S.C., Kohler, P., Kessler, H., Bacher, A. and Richter, G. (2002) Transcriptional regulation by antitermination. Interaction of RNA with NusB protein and NusB/NusE protein complex of *Escherichia coli*. *J. Mol. Biol.*, **316**, 875–885.
- 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 λ . *J. Biol. Chem.*, **267**, 19418–19426.
- Greive, S.J., Lins, A.F. and von Hippel, P.H. (2005) Assembly of an RNA-protein complex. Binding of NusB and NusE (S10) proteins to *boxA* RNA nucleates the formation of the antitermination complex involved in controlling rRNA transcription in *Escherichia coli*. *J. Biol. Chem.*, **280**, 36397–36408.
- Patterson, T.A., Zhang, Z., Baker, T., Johnson, L.L., Friedman, D.I. and Court, D.L. (1994) Bacteriophage lambda N-dependent transcription antitermination. Competition for an RNA site may regulate antitermination. *J. Mol. Biol.*, **236**, 217–228.
- Mason, S.W. and Greenblatt, J. (1991) Assembly of transcription elongation complexes containing the N protein of phage λ and the *Escherichia coli* elongation factors NusA, NusB, NusG, and S10. *Genes Dev.*, **5**, 1504–1512.
- Mizushima, S. and Nomura, M. (1970) Assembly mapping of 30S ribosomal proteins from *E. coli*. *Nature*, **226**, 1214.
- Wimberly, B.T., Brodersen, D.E., Clemons, W.M.J., Morgan-Warren, R.J., Carter, A.P., Vornrhein, C., Hartsch, T. and Ramakrishnan, V. (2000) Structure of the 30S ribosomal subunit. *Nature*, **407**, 327–339.
- Schluenzen, F., Tocilj, A., Zarivach, R., Harms, J., Gluehmann, M., Janell, D., Bashan, A., Bartels, H., Agmon, I., Franceschi, F. et al. (2000) Structure of functionally activated small ribosomal subunit at 3.3 Å resolution. *Cell*, **102**, 615–623.
- Li, S.C., Squires, C.L. and Squires, C. (1984) Antitermination of *E. coli* rRNA transcription is caused by a control region segment containing lambda nut-like sequences. *Cell*, **38**, 851–860.
- Quan, S., Zhang, N., French, S. and Squires, C.L. (2005) Transcriptional polarity in rRNA operons of *Escherichia coli* nusA and nusB mutant strains. *J. Bacteriol.*, **187**, 1632–1638.
- Torres, M., Condon, C., Balada, J.M., Squires, C. and Squires, C.L. (2001) Ribosomal protein S4 is a transcription factor with properties remarkably similar to NusA, a protein involved in both non-ribosomal and ribosomal RNA antitermination. *EMBO J.*, **20**, 3811–3820.
- Torres, M., Balada, J.M., Zellars, M., Squires, C. and Squires, C.L. (2004) *In vivo* effect of NusB and NusG on rRNA transcription antitermination. *J. Bacteriol.*, **186**, 1304–1310.
- Berg, K.L., Squires, C. and Squires, C.L. (1989) Ribosomal RNA operon anti-termination. Function of leader and spacer region *box B-box A* sequences and their conservation in diverse micro-organisms. *J. Mol. Biol.*, **209**, 345–358.
- Luo, X., Hsiao, H.H., Bubunenko, M., Weber, G., Court, D.L., Gottesman, M.E., Urlaub, H. and Wahl, M.C. (2008) Structural and functional analysis of the *E. coli* NusB-S10 transcription antitermination complex. *Mol. Cell*, **32**, 791–802.
- 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. USA*, **91**, 8660–8664.
- Olson, E.R., Tomich, C.S. and Friedman, D.I. (1984) The nusA recognition site: Alteration in its sequence or position relative to upstream translation interferes with the action of the N antitermination function of phage lambda. *J. Mol. Biol.*, **180**, 1053–1063.
- Robledo, R., Gottesman, M.E. and Weisberg, R.A. (1990) λ nutR mutations convert HK022 nun protein from a transcription termination factor to a suppressor of termination. *J. Mol. Biol.*, **212**, 635–643.
- Baron, J. and Weisberg, R.A. (1992) Mutations of the phage lambda nutL region that prevent the action of nun, a site-specific transcription termination factor. *J. Bacteriol.*, **174**, 1983–1989.
- Court, D.L., Patterson, T.A., Baker, T., Costantino, N., Mao, X. and Friedman, D.I. (1995) Structural and functional analyses of the transcription-translation proteins NusB and NusE. *J. Bacteriol.*, **177**, 2589–2591.
- Ward, D.F., DeLong, A. and Gottesman, M.E. (1983) *Escherichia coli* nusB mutations that suppress nusA1 exhibit lambda N specificity. *J. Mol. Biol.*, **168**, 73–85.
- Studier, F.W. (2005) Protein production by auto-induction in high density shaking cultures. *Protein Expr. Purif.*, **41**, 207–234.
- Kabsch, W. (1993) Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. Appl. Crystallogr.*, **26**, 795–800.
- Emsley, P. and Cowtan, K. (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr. D*, **60**, 2126–2132.
- Winn, M.D., Musshudov, G.N. and Papiz, M.Z. (2003) Macromolecular TLS refinement in REFMAC at moderate resolutions. *Methods Enzymol.*, **374**, 300–321.
- Santoro, M.M. and Bolen, D.W. (1988) Unfolding free energy changes determined by the linear extrapolation method. 1. Unfolding of phenylmethanesulfonyl α -chymotrypsin using different denaturants. *Biochemistry*, **27**, 8063–8068.
- Swint, L. and Robertson, A.D. (1993) Thermodynamics of unfolding for turkey ovomucoid third domain: Thermal and chemical denaturation. *Protein Sci.*, **2**, 2037–2049.
- Mayr, L.M., Landt, O., Hahn, U. and Schmid, F.X. (1993) Stability and folding kinetics of ribonuclease T1 are strongly altered by the replacement of cis-proline 39 with alanine. *J. Mol. Biol.*, **231**, 897–912.
- Krissinel, E. and Henrick, K. (2004) Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr. D*, **60**, 2256–2268.
- Altieri, A.S., Mazzulla, M.J., Horita, D.A., Coats, R.H., Wingfield, P.T., Das, A., Court, D.L. and Byrd, R.A. (2000) The structure of the transcriptional antiterminator NusB from *Escherichia coli*. *Nat. Struct. Biol.*, **7**, 470–474.
- Davis, I.W., Murray, L.W., Richardson, J.S. and Richardson, D.C. (2004) MOLPROBITY: Structure validation and all-atom contact analysis for nucleic acids and their complexes. *Nucleic Acids Res.*, **32**, W615–W619.

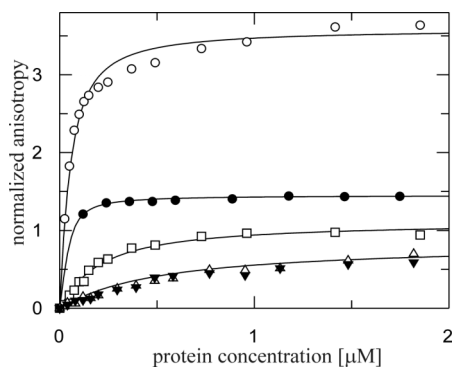
Supplemental Table:

Table S1: Obtained affinity constants for NusB^{D118N} (NusB101) towards different concentrations of fluorescein labeled *rrnG BoxA*. The used buffer conditions were 25 mM HEPES, 100 mM potassiumacetate, pH 7.5.

<u><i>rrnG BoxA</i> [nM]</u>	<u>K_d [nM]</u>
5	66 ± 20
50	50 ± 15
100	100 ± 24

Supplemental Figure Legend:

Figure S1: Fluorescence anisotropy measurements with fluorescein labeled λ *mut* and *rrnG* RNAs towards NusB/NusE ^{Δ loop}. Each titration was performed with 50 nM fluorescein labeled RNA. *rrnG BoxA* (open circles), *rrnG BoxA-spacer* (filled circles), *nutR BoxA* (open squares), *nutR BoxA-spacer* (filled triangles), *nutL BoxA-spacer* (open triangles). Lines represent the best fit to equation (3).

Supplemental Figure:**Fig. S1:**

Einzelarbeit E

Björn M. Burmann, Ulrich Scheckenhofer, Kristian Schweimer und Paul Rösch:

Domain interactions of the transcription:translation coupling factor *E. coli* NusG are intermolecular and transient.. *Eingereicht.*

Domain interactions of the transcription:translation coupling factor *E. coli* NusG are intermolecular and transient.

Björn M. Burmann, Ulrich Scheckenhofer, Kristian Schweimer, and Paul Rösch*

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

Running title: NusG domain interactions

* 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 bacterial transcription factor NusG is suggested to act as a key coupling factor between transcription and translation (1) and contributes to Phage λ mediated antitermination in *E. coli* that enables read-through of early transcription termination sites. *E. coli* NusG consists of two structurally and functionally distinct domains that are connected *via* a flexible linker. The homologous *Aquifex aeolicus* NusG with a secondary structure that is highly similar to *E. coli* NusG shows, under certain conditions, direct interaction between its N-terminal and its C-terminal domain resulting in a domain-swapped dimer. Here we performed NMR paramagnetic relaxation enhancement measurements and identified interdomain interactions that were, however, concentration dependent and thus considered not only weak and transient, but also predominantly intermolecular, and this notion of two virtually independent domains was supported by ^{15}N relaxation measurements. Thus, a regulatory role of NusG interdomain interactions is considered unlikely.

E. coli NusG (N-utilization substance G) is essential for cell viability (2) and found in all known bacteria. The C-terminal domain, CTD, contains a KOW motif that is also found in archaeal (3) and eukaryotic (4) proteins. NusG's exact role

in transcription regulation remained obscure until very recently when NusG was identified as the long sought after coupling factor between transcription and translation in *E. coli* (1), in addition to its better known role in the transcription-elongation complex (TEC) where it increases the elongation rate of the RNAP *in vivo* and *in vitro* (5-7) by suppression of transcriptional pausing (7). NusG is also a component of the phage λ protein N mediated antitermination (8-10) and the phage HK022 protein Nun mediated termination complex (11, 12), along with the other Nus factors A, B, E (ribosomal protein S10), RNA, and the RNAP. The regulatory functionality of NusG can be attributed to direct interaction with the RNAP (8, 13, 14).

Additionally, NusG is involved in transcription antitermination in ribosomal RNA (rRNA)(15, 16) and at certain ρ dependent sites (11, 17, 18). NusG recruits ρ and prevents backtracking of the RNAP (7, 18), but detailed information on the final step of ρ -dependent termination is still lacking (19). Direct NusG: ρ interaction was reported (17, 18), and the NusG carboxy-terminal domain (CTD) was identified as the interaction domain (1). In addition to its role in transcription, NusG is also involved in translational regulation (20), and the NusG CTD was again identified as the ribosome interaction domain (1).

The structure of *E. coli* NusG (19) is highly similar to the structures of the homologous *Aquifex aeolicus* (21, 22) and *Thermus thermophilus* (23) proteins. The *A. aeolicus* NusG is reported to form a domain-swapped dimer under certain crystallization conditions (22), although this observation is disputed as such a dimer is not observed in a different crystallization setup (21) and related studies (19, 24). Interestingly, however, the NusG paralogue RfaH is reported to exist in a closed conformation with both domains tightly interacting, and this tight interaction is abolished when the RfaH-CTD leaves the binding pocket on the RfaH-NTD upon binding to the *ops*-site of the non-template DNA strand, thus providing an RfaH activation step (24, 25). Although *E. coli* NusG and RfaH exhibit the same folding topology for their NTDs (19, 25) (NusG-NTD, PDB-ID: 2K06; RfaH, PDB-ID: 2OUG), their CTDs show reversed topologies: Whereas RfaH-CTD is all α -helical, NusG-CTD exhibits an anti-parallel β -barrel-like structure (19) (NusG-CTD, PDB-ID: 2JVV).

The area masked on the RfaH-NTD by its CTD consists of several surface exposed hydrophobic residues that are directly involved in RfaH:RNAP interaction (14, 26). Surface hydrophobic residues are more numerous in RfaH than in NusG (19) where they are supposed to be involved in additional protein interactions (21).

Here we study individual NusG domains to gain insight into their role within the TEC, their relative motional behavior and their mutual interactions by chemical shift and paramagnetic relaxation enhancement (PRE) perturbation titrations as well as by relaxation experiments.

Materials and Methods

Cloning, expression, and purification of full-length NusG and the individual domains: Cloning, expression and purification was based on published methods (18, 19). NusG was cloned *via* Bpu1102I and NdeI restriction sites into the *E. coli* expression vector pET11A

(Novagen, Madison, WI, USA). *E. coli* strain BL21(DE3) (Novagen, Madison, WI, USA) harboring the recombinant plasmid was grown at 37 °C in LB medium (Luria-Bertani) containing ampicillin (100 μ g/ml) until an $OD_{600} = 0.8$ was reached and the cells were induced to a final concentration of 1 mM isopropyl 1-thio- β -D-galactopyranoside (IPTG). Cells were harvested 4 h after induction, resuspended in 4 times the pellet weight of lysis buffer (50 mM TRIS, 150 mM NaCl, 1 mM DTT, pH 7.5, 1 Protease inhibitor tablet (Complete, EDTA-free, Roche Diagnostics GmbH, Mannheim, Germany)), and lysed by using a micro-fluidizer (Microfluidics, Newton, MA, USA). After centrifugation Polyethylenimine (Fluka, Basel, Switzerland) was added drop-wise under continuous stirring to the supernatant to a final concentration of 0.6%. The lysate was incubated for 20 min and centrifuged at 12.000 g. Ammonium sulfate was added drop-wise to a final concentration of 60% under stirring to the supernatant. The lysate was centrifuged at 12.000 g and the pellet was solved in 30 ml of buffer A (50 mM TRIS, pH 7.5). The lysate was afterwards dialyzed against 2 x 4 l overnight against buffer A. The lysate was applied to a HeparinFF column (GE Healthcare, Munich, Germany) using a step gradient with increasing NaCl concentrations (0 – 1 M). For further purification the eluted fractions containing NusG were pooled and concentrated with Vivaspin concentrators (Vivascience, MWCO 5,000 Da). The concentrated sample was applied to a S75 gel filtration column (50 mM TRIS, 150 mM NaCl; GE Healthcare, Munich, Germany). The fractions containing NusG were pooled and dialyzed against buffer as used for NMR measurements (10 mM potassium phosphate, 50 mM NaCl, pH 6.4). The identity and structural integrity of purified protein was analyzed by SDS- PAGE as well as by NMR spectroscopy. NusG-NTD (1-124) was cloned *via* BPU1102I and NdeI into the *E. coli* expression vector pET11A (Novagen, Madison, WI, USA). The same procedure as described above for full-length NusG was performed for expression and purification of NusG-NTD.

NusG-CTD (123-181) was cloned and purified as described before (1).

Full-length NusG mutations S16C and S163C: For NusG^{S16C} the following primers were used: 5'-primer: GTC GTT CAG GCG TTT TGC GGT TTT GAA GGC CGC, 3'-primer: GCG GCC TTC AAA ACC GCA AAA CGC CTG AAC GAC. For NusG^{S163C} the following primers were used: 5'-primer: CTG AAA GTG TCT GTT TGT ATC TTC GGT CGT GCG, 3'-primer: CGC ACG ACC GAA GAT ACA AAC AGA CAC TTT CAG (Purimex, Grebenstein, Germany). Mutations were introduced by using the QuikChange protocol (Stratagene, La Jolla, CA, USA). Expression and purification were the same described before for NusG full-length besides the addition of 1 mM DTT to all buffers used.

Random spin labeling of the ϵ -amino groups of lysines: Spin labeling of the lysines with 1-Oxyl-2,2,5,5-tetramethylpyrroline-3-carboxylate N-hydroxysuccinimide Ester (OXYL-1-NHS; Toronto research Chemicals Inc., North York, ON, Canada) was done by a slightly modified published protocol (27). Briefly, 500 μ M protein solution in NMR buffer (10 mM potassium phosphate, 50 mM NaCl, pH 6.4) were washed with 10 mM sodium carbonate buffer, pH 9.2, for buffer exchange in Vivaspin concentrators (Vivascience, MWCO 5,000 Da). A 10 fold excess of OXYL-1-NHS in DMSO was added, followed by incubation for one hour at room temperature and an additional hour at 4 °C in the dark. To remove unreacted OXYL-1-NHS, the sample was washed with 20 ml of NMR-buffer in a Vivaspin concentrator. For reduction of the spin label, ascorbate (500 mM stock) was added directly to the NMR tube to a final concentration of 5 mM.

Site specific spin labeling of NusG cysteine mutants: Spin labeling of the cysteines introduced into NusG, NusG^{S16C} and NusG^{S163C}, respectively, with (1-Oxyl-2,2,5,5-tetramethyl- Δ 3-pyrroline-3-methyl) Methanethiosulfonate (MTSL; Toronto research Chemicals Inc., North

York, ON, Canada) was done by a slightly modified published protocol (28). Briefly, 500 μ M protein solution in NMR buffer (10 mM potassium phosphate, 50 mM NaCl, pH 6.4) were washed with 10 mM acetate buffer for buffer exchange in Vivaspin concentrators (Vivascience, MWCO 5,000 Da). After addition of DTT to a final concentration of 5 mM, the sample was kept for one hour at 4 °C. For removal of DTT, the solution was eluted isocratically with 10 mM acetate from a HiTrap Desalting column (GE Healthcare, Munich). A 10 fold excess of MTSL dissolved in acetonitrile was added, followed by incubation overnight at room temperature in the dark. To remove unreacted MTSL, the sample was again washed with 20 ml of NMR-buffer in a Vivaspin concentrator. For reduction of the spin label, ascorbate (500 mM stock) was added directly to the NMR tube to a final concentration of 5 mM.

NMR: NMR experiments were recorded at sample temperatures of 298 K on Bruker Avance 600 MHz and 700 MHz (equipped with a cryogenically cooled probe) NMR spectrometers.

Assignments for NusG-NTD and full-length NusG were from previous work (19). The isolated NusG-CTD ¹H,¹⁵N-HSQC resonances matched the corresponding signals of full-length NusG perfectly and were trivially assigned and verified by triple resonance NMR experiments.

Paramagnetic relaxation enhancements ($R_{2,para} = R_{2,spinlabel} - R_{2,no\ spinlabel}$) were determined by a two point scheme using an HSQC experiment with additional spin echo period during the first INEPT transfer (29). Spin echo intervals were set to 0.1 ms and 10.2 ms. ¹⁵N longitudinal (R_1) and transverse (R_2) relaxation rates were determined by standard methods at a ¹H frequency of 600.2 MHz.

Results and Discussion

NMR titrations suggest independence of the NusG domains: ¹H, ¹⁵N chemical shifts

of resonances from isolated NusG-CTD and isolated NusG-NTD as observed in HSQC (heteronuclear single quantum coherence) spectra were virtually identical to the chemical shifts of the respective domains in full-length NusG (Fig. S1), rendering persistent domain interactions within full-length NusG highly unlikely. In order to further clarify whether or not transient domain interactions exist for the more general case of the isolated domains we resorted to observation of HSQC chemical shift perturbations of an isotope labeled domain, NusG-CTD or NusG-NTD, upon addition of the unlabeled potential interacting domain, a method that is well established as a tool to study the interplay between molecules (30). NMR spectroscopy, due to its inherent insensitivity and its resulting requirement for sample concentrations in the high micromolar range, is useful for detecting even weak ($k_d \sim \mu\text{M}$ – mM) interactions, and perturbations of amide group resonance shifts as detected by ^1H , ^{15}N HSQC are very sensitive even to subtle structural changes. As NusG-CTD F165 was claimed to strongly interact with a NusG-NTD hydrophobic cavity (22), significant chemical shift changes are expected on domain contact at least for F165 and residues in its vicinity, that is residues in the loop between strands $\beta 3$ and $\beta 4$. Titration of unlabeled NusG-NTD to ^{15}N labeled NusG-CTD to a twofold excess and *vice versa*, however, did not result in observable chemical shift changes in the NMR experiments (Fig. S2). Thus, with the level of sensitivity provided by HSQC perturbation experiments, domain interaction could be detected neither within the full-length protein nor for the isolated domains under our experimental conditions. This observation strongly correlates with our earlier conclusion that the area around F165 is not involved in mutual interactions of NusG domains, but is rather a key residue in NusE:NusG complex formation (1).

^{15}N relaxation reveals decoupled domain reorientation: In order to study the degree of motional independence of the NusG domains we determined their ^{15}N relaxation rates in the context of the full-

length protein (31). In a two-domain protein, concerted tumbling of domains can be described by a single rotational diffusion tensor, whereas independent tumbling of domains requires description with different rotational diffusion tensors (32). In an isotropic rotation model, differences in these tensors are directly reflected in differences in the average relaxation rates.

^{15}N transverse and longitudinal relaxation rates were determined at 14.1 T with a sample concentration of 200 μM to reduce aggregation (Fig. 1A, 1B). In the HSQC spectrum of the full-length protein, virtually all residues of NusG-CTD were observed, the average longitudinal relaxation rate (R_1) was $1.35 (\pm 0.09) \text{ sec}^{-1}$ and the transversal relaxation rate (R_2) was $15.4 (\pm 1.9) \text{ sec}^{-1}$, corresponding to a rotational correlation time $\tau_c = 10.1 \text{ ns}$ in an isotropic model. Although the NusG-NTD signals were considerably weaker, 67 non-overlapping signals were observed and the average longitudinal relaxation rate was $1.1 (\pm 0.2) \text{ sec}^{-1}$, the transversal rate $19.0 (\pm 4.6) \text{ sec}^{-1}$, corresponding to $\tau_c = 13.1 \text{ ns}$ in the isotropic model. Analysis of R_1/R_2 distributions offers an elegant method to detect interdomain motion on the time scale faster than the overall tumbling (33). The R_1/R_2 ratios form the basis for determination of the rotational diffusion tensor by NMR relaxation, and for compact globular proteins a uniform distribution of R_1/R_2 ratios is characteristic (31). The bimodal distribution of the R_1/R_2 ratios (Fig. 1C) that reflect the two domains demonstrates their different rotational reorientation behavior, arguing against a stable domain interaction on the ns timescale. Although the difference in relaxation rates demonstrates independent movement of the two domains to a certain degree, τ_c of the CTD in full length NusG is higher than τ_c of the isolated domain (10.1 ns vs. 4 ns). This indicates that motional decoupling of NusG-CTD and NusG-NTD *via* the 5 residue linker is imperfect, and a significant contribution of the overall rotation of the full length NusG to the relaxation mechanisms is expected. In intact NusG, V162 and I164 of NusG-CTD exhibit significantly enhanced R_2

rates of 21.5 sec^{-1} as compared to the domain average of 15.4 sec^{-1} , a difference not found in isolated CTD. Most likely the enhanced transverse relaxation can be attributed to a chemical exchange contribution, and as these residues are located in the domain interface of the swapped-dimer crystal structure, this might be regarded as initial evidence of transient domain interaction in solution.

Only highly sensitive PRE experiments show signs of interaction between the isolated domains: Fast exchange on the NMR timescale leads to observation of population averaged NMR parameters, and states that are extremely weakly populated on time average often cannot be detected even by HSQC chemical shift perturbations or changes in relaxation rates. Paramagnetic interactions, however, provide a means to detect even weak and transient interactions between molecules. Paramagnetic centers dramatically increase relaxation rates of nearby nuclei and, as a consequence, presence of even very minor concentrations of paramagnetic labels in proximity to observed nuclei can enhance the relaxation rate of the latter to an observable degree (34). Thus, PRE observed in a non spin-labeled protein in the presence of a spin-labeled protein points to an at least transient proximity of both molecules. The experiment is straightforward as a non-spin-labeled protein, for example NusG-CTD, can be made easily detectable by ^{15}N enrichment in HSQC spectra, and introduction of paramagnetic centers into proteins, for example NusG-NTD, is possible by random labeling of mostly surface exposed lysine residues with 1-Oxyl-2,2,5,5-tetramethylpyrroline-3-carboxylate N-hydroxysuccinimide Ester (OXYL-1-NHS) (35). In such an experiment, addition of OXYL-1-NHS-NusG-NTD to ^{15}N -NusG-CTD causes observable and specific PRE in the latter, and *vice versa* (Fig. 2). Mapping of the resonances with increased relaxation rates onto the three-dimensional structures of NusG-CTD shows very clearly the region around F165 and the region P140 to N145, close to F165 in the turn between β^1 and β^2 , to be the ones most seriously affected (Fig. 2).

The outcome of the inverse experiment, that is addition of OXYL-1-NHS-NusG-CTD to ^{15}N -NusG-NTD, was not so clear-cut. The hydrophobic patch around F65 was clearly affected, together with several residues in helix α^3 and residues close to F65 that form an apolar surface surrounded by polar residues. However, several isolated residues in helices α^1 and α^2 at the opposite side of the molecule as well as several residues of the »upper« loop regions facing away from the putative interaction side were affected (Fig. 2). Combined, the PRE results yield a picture in which a plug around F165 fits nicely into a socket around F65. Although this is a reasonable contact surface, and although these sites correspond well to those found in the swapped dimer crystal structure (22), our negative results with HSQC chemical shift mapping indicate the interaction to be rather weak and transient.

PREs of site-specific full-length NusG are concentration dependent: Spin-labeling of cysteines with 1-Oxyl-2,2,5,5-tetramethyl- Δ^3 -pyrroline-3-methyl)

Methanethiosulfonate (MTSL) provides a means to observe the PRE induced by site-specific labels. Thus, in order to clarify whether or not the weak domain interaction observed by random PRE in solution is intra- or intermolecular, we performed additional PRE experiments on two full-length NusG constructs, one containing an MTSL-cysteine at position S16C, the other containing an MTSL-cysteine at position S163C. For both mutants, PRE of the spin-label on the respective other domain was observed, clearly indicating an interaction between NusG-NTD and NusG-CTD and thus neither supporting nor ruling out the formation of a domain-swapped dimer. However, intramolecular domain interaction would result in a concentration-independent PRE, whereas concentration dependent PRE would demonstrate intermolecular domain interactions. PRE determined at different protein concentrations (Fig. 3) clearly show a direct dependence of PRE on sample concentration, thus ruling out an interaction within one molecule and

favoring the effect related to the transient formation of a domain-swapped dimer.

Conclusions: Our NMR studies on domain interactions of *E. coli* NusG show these interactions to be not detectable *via* classical chemical shift mapping, and ^{15}N relaxation measurements reveal independent movements of the two domains to a very high degree. A domain interaction could only be observed with the extremely sensitive technique of PRE measurements, and it needed a combination of this technique and site-directed paramagnetic labeling to show

that these weak interactions were intermolecular rather than intramolecular. Thus, a low population of molecules seems to adopt a dimolecular conformation compatible with the proposed swapped dimer observed in X-ray crystallography. Intramolecular domain interaction, however, would be required for an autoinhibitory function to reduce the population of the active open state. The mode of regulation found for the paralog RfaH (25) can thus be ruled out to work for NusG, which may explain why NusG maintains its overall function, while RfaH needs the *ops* site to render a functional protein.

REFERENCES

1. Burmann, B. M., Schweimer, K., Luo, X., Wahl, M. C., Stitt, B. L., Gottesman, M. E., and Rösch, P. (in press) *Science*
2. Downing, W. L., Sullivan, S. L., Gottesman, M. E., and Dennis, P. P. (1990) *J. Bacteriol.* **172**, 1621-1627
3. Kyrpides, N. C., and Ouzounis, C. A. (1999) *Proc. Natl. Acad. Sci. USA* **96**, 8545-8550
4. Hartzog, G. A., Wada, T., Handa, H., and Winston, F. (1998) *Genes Dev.* **12**, 357-369
5. Burova, E., Hung, S., Sagitov, V., Stitt, B., and Gottesman, M. (1995) *J. Bacteriol.* **177**, 1388-1392
6. Burns, C. M., Richardson, L. V., and Richardson, J. P. (1998) *J. Mol. Biol.* **278**, 307-16
7. Artsimovitch, I., and Landick, R. (2000) *Proc. Natl. Acad. Sci. USA* **97**, 7090-7095
8. Li, J., Horwitz, R., McCracken, S., and Greenblatt, J. (1992) *J. Biol. Chem.* **267**, 6012-9
9. Sullivan, S. L., Ward, D. F., and Gottesman, M. E. (1992) *J. Bacteriol.* **174**, 1339-44
10. Mason, S. W., Li, J., and Greenblatt, J. (1992) *J. Biol. Chem.* **267**, 19418-19426
11. Sullivan, S. L., and Gottesman, M. E. (1992) *Cell* **68**, 989-994
12. Burova, E., Hung, S. C., Chen, J., Court, D. L., Zhou, J. G., Mogilnitskiy, G., and Gottesman, M. E. (1999) *Mol. Microbiol.* **31**, 1783-93
13. Mason, S. W., and Greenblatt, J. (1991) *Genes Dev.* **5**, 1504-1512
14. Belogurov, G. A., Mooney, R. A., Svetlov, V., Landick, R., and Artsimovitch, I. (2009) *EMBO J.* **28**, 112-122
15. Torres, M., Balada, J. M., Zellars, M., Squires, C., and Squires, C. L. (2004) *J. Bacteriol.* **186**, 1304-10
16. Zhou, Y., Filter, J. J., Court, D. L., Gottesman, M. E., and Friedman, D. I. (2002) *J. Bacteriol.* **184**, 3416-8
17. Li, J., Mason, S. W., and Greenblatt, J. (1993) *Genes Dev.* **7**, 161-172
18. Pasman, Z., and von Hippel, P. H. (2000) *Biochemistry* **39**, 5573-5585
19. Mooney, R. A., Schweimer, K., Rösch, P., Gottesman, M. E., and Landick, R. (2009) *J. Mol. Biol.* **391**, 341-358
20. Zellars, M., and Squires, C. L. (1999) *Mol. Microbiol.* **32**, 1296-304
21. Steiner, T., Kaiser, J. T., Marinkovic, S., Huber, R., and Wahl, M. C. (2002) *EMBO J.* **21**, 4641-4653

22. Knowlton, J. R., Bubunenko, M., Andrykovitch, M., Guo, W., Routzahn, K. M., Waugh, D. S., Court, D. L., and Ji, X. (2003) *Biochemistry* **42**, 2275-2281
23. Reay, P., Yamasaki, K., Terada, T., Kuramitsu, S., Shirouzu, M., and Yokoyama, S. (2004) *Proteins* **56**, 40-51
24. Artsimovitch, I., and Landick, R. (2002) *Cell* **109**, 193-203
25. Belogurov, G. A., Vassilyeva, M. N., Svetlov, V., Klyuyev, S., Grishin, N. V., Vassilyev, D. G., and Artsimovitch, I. (2007) *Mol. Cell* **26**, 117-129
26. Sevostyanova, A., Svetlov, V., Vassilyev, D. G., and Artsimovitch, I. (2008) *Proc. Nat. Acad. Sci. USA* **105**, 865-870
27. Lawrence, J. J., Berne, L., Ouvrier-Buffet, J. L., and Piette, L. H. (1980) *Eur. J. Biochem.* **107**, 263-269
28. Berliner, L. J., Grunwald, J., Hankovszky, H. O., and Hideg, K. (1982) *Anal. Biochem.* **119**, 450-5
29. Iwahara, J., Tang, C., and Marius Clore, G. (2007) *J. Magn. Reson.* **184**, 185-195
30. Zuiderweg, E. R. (2002) *Biochemistry* **41**, 1-7
31. Kay, L. E., Torchia, D. A., and Bax, A. (1989) *Biochemistry* **28**, 8972-8979
32. Fushman, D., Varadan, R., Assfalg, M., and Walker, O. (2004) *Prog. NMR Spect.* **44**, 189-214
33. Horstmann, M., Ehses, P., Schweimer, K., Steinert, M., Kamphausen, T., Fischer, G., Hacker, J., Rösch, P., and Faber, C. (2006) *Biochemistry* **45**, 12303-12311
34. Tang, C., Iwahara, J., and Clore, G. M. (2006) *Nature* **444**, 383-386
35. Jahnke, W., Rudisser, S., and Zurini, M. (2001) *J. Am. Chem. Soc.* **123**, 3149-3150

ACKNOWLEDGMENT

We thank Ramona Heissmann for excellent technical assistance.

FOOTNOTES

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

Figure Legends:

Fig. 1 A) Longitudinal relaxation rates (R_1) for full-length NusG . B) Transversal relaxation rates (R_2) for full-length NusG . Residues 45 to 65 and 115 to 125 are highly flexible and located in strongly overlapping regions of the protein and their relaxation rates could not be determined. C) The distribution of R_1/R_2 is bimodal for full-length NusG. Bars representing NusG-CTD are shown in black and bars representing NusG-NTD are shown in gray.

Fig. 2 A) Surface representation in gray of the individual *E. coli* NusG domains (NusG-NTD PDB-ID: 2K06; NusG-CTD PDB-ID: 2JVV; (19)) with secondary structure elements. Central F65 and F165 are shown as sticks in blue, aminoacids with an $R_{2,para}(H^N)$ -effect >20 Hz in the titrations of the individual domains are highlighted in red. B) $R_{2,para}(H^N)$ -rates for each aminoacid upon titration with the individual spin labeled domains. Values and their respective deviations are in gray. Signals disappearing due to extensive line broadening are highlighted in red. The dotted line represents the significance level of 20 Hz.

Fig. 3 Effects of spin labeled NusG^{S16C} on the $R_{2,para}(H^N)$ -rates of NusG-CTD in full-length NusG. Dark gray bars: effects at a protein concentration of 150 μ M for selected aminoacids. Light gray bars: effects at a protein concentration of 50 μ M for these aminoacids.

Figure 1

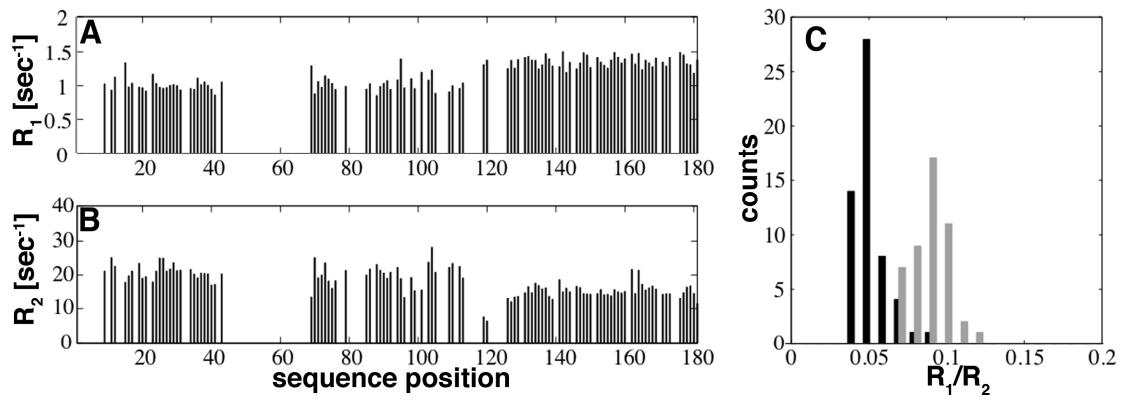
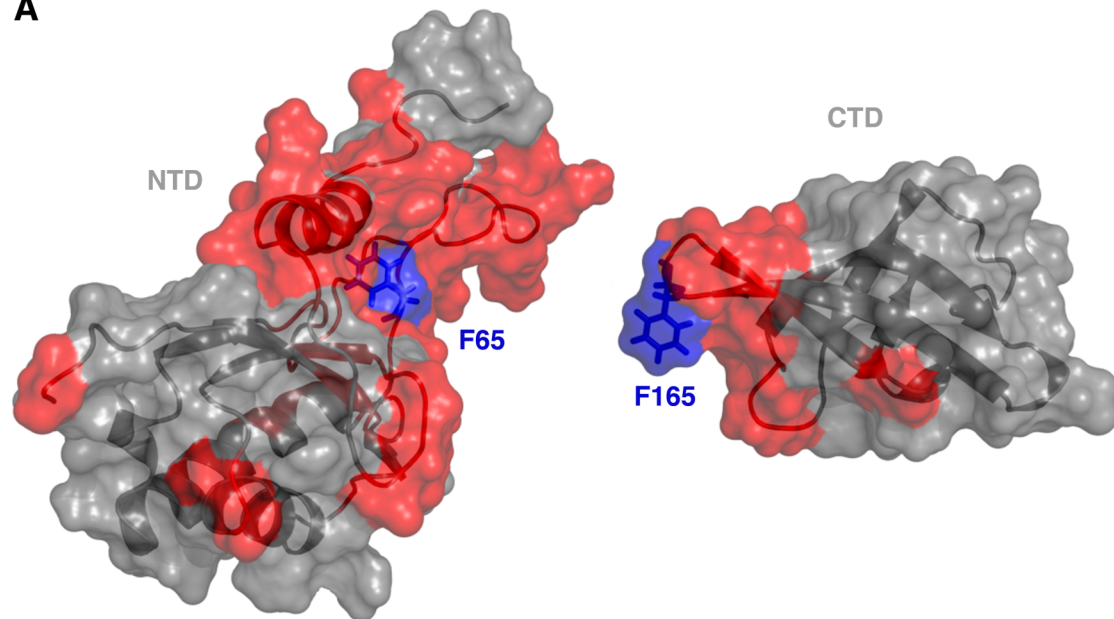


Figure 2

A



B

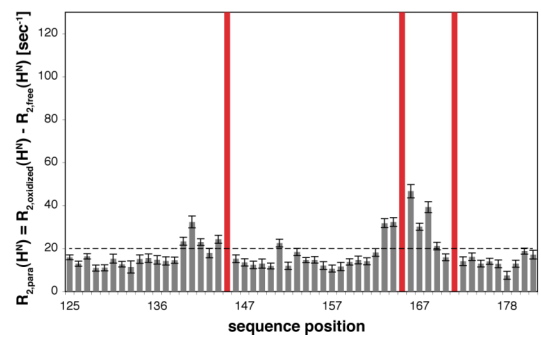
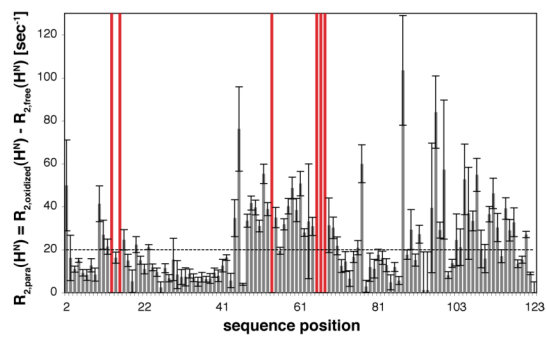
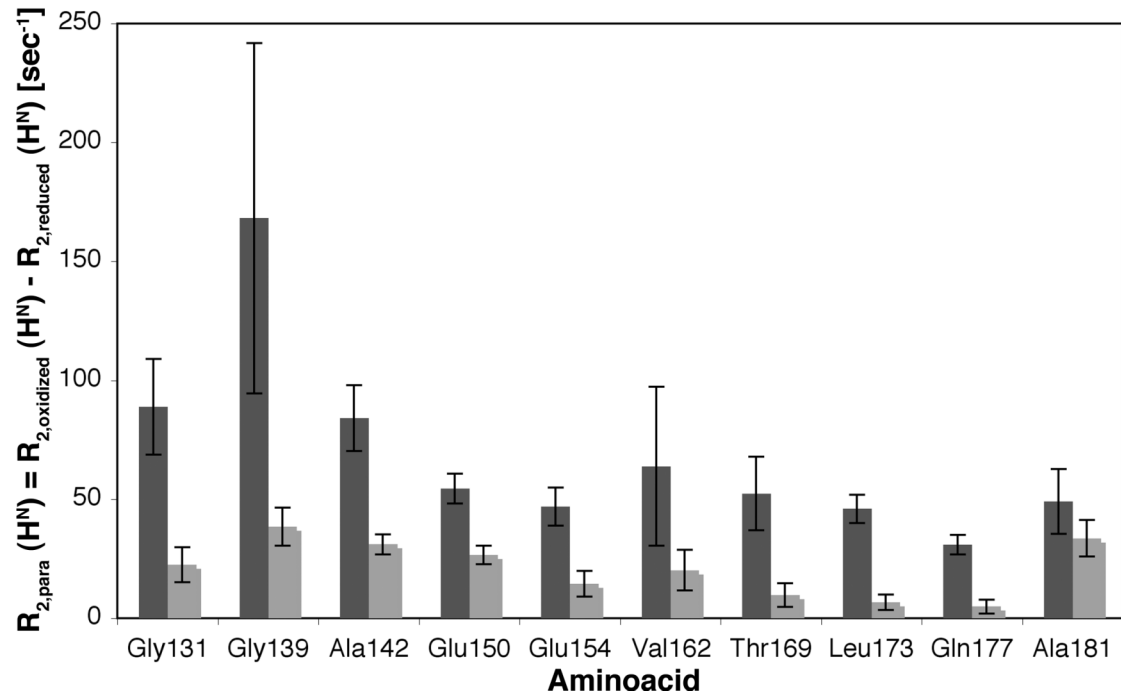


Figure 3



Supplemental Figure Legends:

Fig. S1: Overlay of the ^1H , ^{15}N HSQC spectra of full-length NusG (black), A) the individual NusG-NTD (red), and B) the individual NusG-CTD (red).

Fig S2: ^1H , ^{15}N HSQC titrations of the individual NusG domains. A) Overlay of ^{15}N -NusG-NTD (black) titrated with unlabeled NusG-CTD to a 1:2 complex (red). B) Overlay of ^{15}N -NusG-CTD (black) titrated with unlabeled NusG-NTD to a 1:2 complex (red). Concentrations in each experiment were about 300 μM .

Figure S1

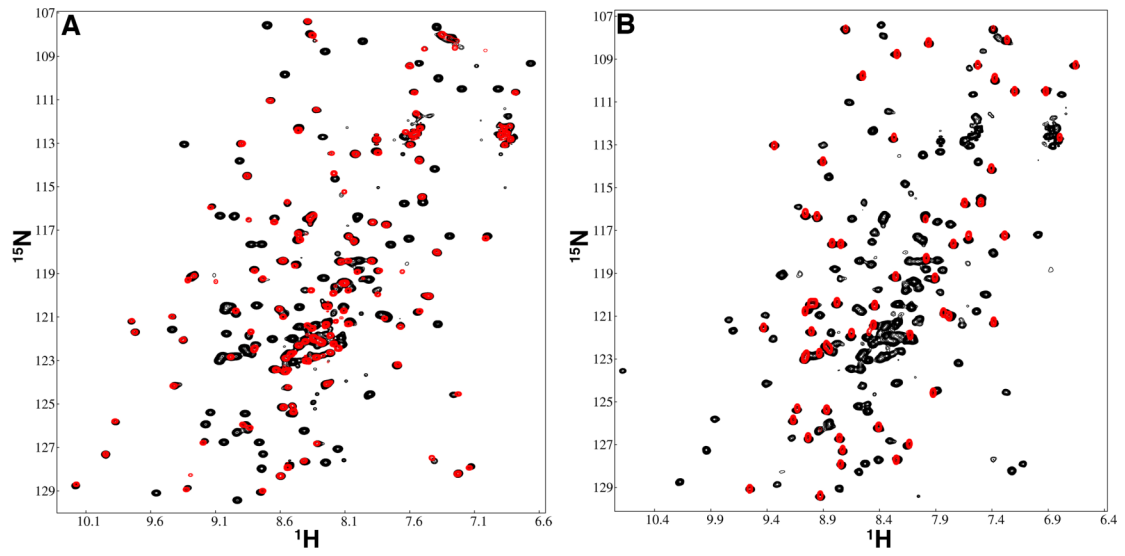
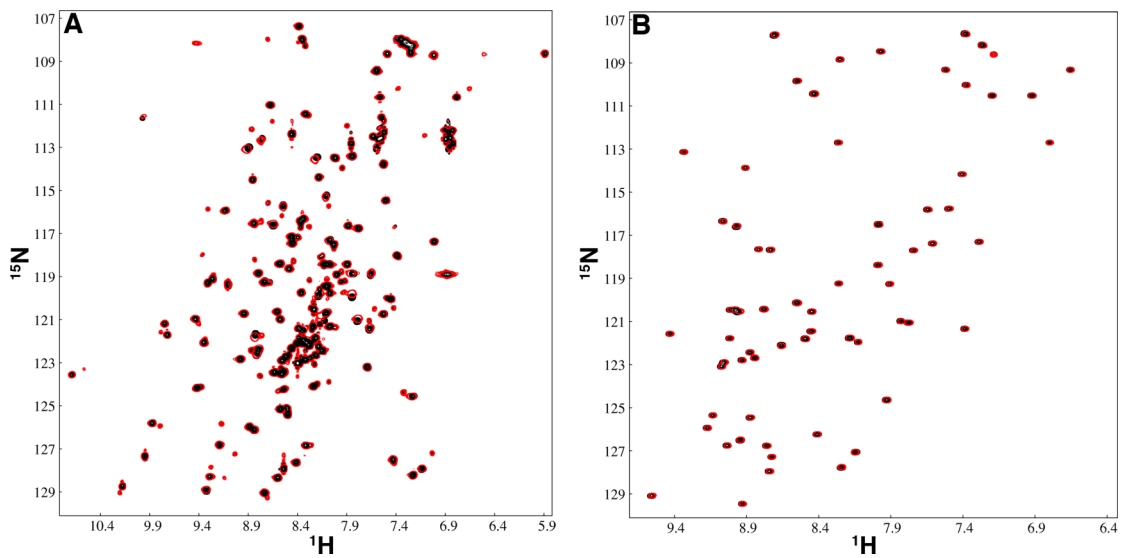


Figure S2:



Einzelarbeit F

Björn M. Burmann, Kristian Schweimer, Xiao Luo, Markus C. Wahl, Barbara L. Stitt, Max E. Gottesman und Paul Rösch:

A NusE:NusG Complex Links Transcription and Translation. *Science*, **328**, 501 – 504, *Authors manuscript* (2010).

A NusE:NusG Complex Links Transcription and Translation

Björn M. Burmann¹, Kristian Schweimer¹, Xiao Luo^{2,3}, Markus C. Wahl^{2,3},
Barbara L. Stitt⁴, Max E. Gottesman⁵, and Paul Rösch¹

¹Lehrstuhl Biopolymere und Forschungszentrum für Bio-Makromoleküle, Universität Bayreuth, Universitätsstraße 30, 95447 Bayreuth, Germany

²Max-Planck-Institut für biophysikalische Chemie, Makromolekulare Röntgenkristallographie, Am Faßberg 11, 37077 Göttingen, Germany

³Freie Universität Berlin, Fachbereich Biologie-Chemie-Pharmazie, Institut für Chemie und Biochemie, AG Strukturbiochemie, Takustr. 6, 14195 Berlin, Germany

⁴Department of Biochemistry and Fels Institute for Cancer Research and Molecular Biology, Temple University School of Medicine, Philadelphia, Pennsylvania 19140

⁵Department of Microbiology and Institute of Cancer Research, Columbia University Medical Center, New York, New York 10032

Bacterial NusG is a highly conserved transcription factor that is required for most Rho activity *in vivo*. We show by nuclear magnetic resonance spectroscopy that *E. coli* NusG carboxy-terminal domain forms a complex alternatively with Rho or with transcription factor NusE, a protein identical to 30S ribosomal protein S10. Since NusG amino-terminal domain contacts RNA polymerase and the NusG carboxy-terminal domain interaction site of NusE is accessible in the ribosomal 30S subunit, NusG may act as a link between transcription and translation. Uncoupling of transcription and translation at the ends of bacterial operons enables transcription termination by Rho factor, and competition between ribosomal NusE and Rho for NusG helps to explain why Rho cannot terminate translated transcripts.

E. coli NusG is a two domain protein (Fig. S1; (1)) that is essential for cell viability (2). NusG homologs are found in all known bacteria, and the 27 amino acid NusG carboxy-terminal domain (CTD) KOW motif is found in proteins from *archaea* and eukaryotes (3-5). Sequence highly homologous to NusG amino-terminal domain (NTD) followed by KOW motifs appears in human transcription factor hSpt5 (6). NusG suppresses RNA polymerase (RNAP) pausing and increases elongation rates *in vitro*. *In vivo*, it acts in concert with NusA, NusB and NusE to promote read-through of terminators within ribosomal *rrn* operons and on the phage λ chromosome, a process that additionally requires the λ N protein (7). NusG activates Rho transcription termination factor *in vitro* and is necessary for most Rho-mediated termination events *in vivo* (8, 9). NusG-NTD binds to RNAP and increases the rate of transcription elongation but cannot stimulate termination (1, 10).

The rates of transcription and translation are correlated over a range of different growth rates (11), and NusG was suggested to be involved in this correlation (12). Thus, depletion of NusG slowed the rate of *lacZ* translation without affecting the rate of *lacZ* transcription elongation (12). The dual capacity of NusG to act in transcription as well as in translation is shared by the 30S ribosomal subunit protein NusE which doubles as a component of some transcription elongation complexes (TECs) (13). As a transcription factor, NusE is loaded by NusB onto the *boxA* sequence within *nut* RNA (14-16) and becomes part of an antitermination complex that includes NusA, NusG, and other cellular factors (7, 17). The NusB:NusE:RNA ternary complex is proposed to associate with RNAP through NusE (7, 18).

Genetic evidence supports an interaction between NusG and NusE. Thus, the *nusG4* (S163F) mutation restores λ N antitermination in a *nusE71* (*nusEA86D*) strain (19). We asked if this genetic interaction reflects a direct physical contact between the proteins. For all experiments, we used the NusE ^{Δ loop} variant (15), referred to here as NusE. NusE ^{Δ loop} is fully active for transcription, although it cannot support translation (Fig. S2; (20)), and its crystal structure is known in the NusB:NusE complex (15). We analyzed mixtures of NusG and the NusB:NusE complex by size exclusion chromatography. A mixture of NusB:NusE and NusG eluted earlier from the column than either NusB:NusE or NusG alone (Fig. S3; (20)), consistent with formation of a NusB:NusE:NusG complex. To confirm the interaction and to map contact surfaces, we investigated complex formation by NMR. Titration of isolated ¹⁵N labeled NusG-NTD or NusG-CTD with NusB:NusE complex caused chemical shift changes in the

^1H , ^{15}N -HSQC (heteronuclear single quantum coherence) NMR spectrum of NusG-CTD, but not of NusG-NTD (Fig. S4, S5). Reverse labeling (^{15}N -NusE or ^{15}N -NusB, unlabeled NusG-CTD) revealed that NusE is the recognition protein in the NusB:NusE complex (Fig. S4, S5), suggesting direct NusG-CTD:NusE interaction.

From the chemical shift changes upon titration, we could estimate the dissociation constant for the NusB:NusE:NusG-CTD (molecular weight, mw: 32.3 kDa) interaction as $K_d = 50 \mu\text{M}$ (Fig. S6). Comparison of secondary chemical shifts and characteristic nuclear Overhauser enhancement spectroscopy (NOESY) cross peak patterns of NusB:NusE and NusG-CTD with the corresponding data of the NusB:NusE:NusG-CTD complex revealed no substantial conformational changes in any of the participating proteins, indicating that only minor side chain rearrangements are necessary to form the interaction surfaces (Fig. 1). Isotope-filtered NOESY spectra (e.g. Fig. S7; (21)) revealed unambiguous intermolecular contacts: NusG-I164:NusE-M88, NusG-I164:NusE-I100, NusG-T169:NusE-Q99, NusG-P140:NusE-V84, NusG-R167:NusE-V98, NusG-E172:NusE-S101(22) (Fig. 2A). Docking of the rigid domains with flexible side chains using the NOESY-derived intermolecular NOEs as distance restraints yielded a distinct conformation of the complex without any violation of the restraints (Fig S8). The NusE-binding region on NusG-CTD is composed of loops between β -strands β 1 and β 2, β 3 and β 4, and residues from β -strand β 4. Hydrophobic amino acids P140, F141, F144, I164, and F165 show close contacts to hydrophobic residues of NusE. NusG-R167 is very close to NusE-D87, and NusG-F165 at the tip of the loop between β -strands β 3 and β 4 is buried deeply in a hydrophobic pocket on NusE (Fig. 2B, C). The NusE hydrophobic pocket is composed of residues from helix α 2 and the carboxy-terminal β -strand β 4. There were no significant effects on resonances of NusB in the NMR spectra of the NusB:NusE complex upon binding of NusG-CTD. These data together with mapping of the chemical shift changes on the sequences and the known three-dimensional structures (Fig. 1) show that the NusE-binding interface is opposite to the NusB:NusE interaction region, in proximity to the ribosome-anchoring flexible loop, R46 to T67 (Fig. S2).

NusE residues M88 and D97 and NusG residues P140, F165, and R167 in the NusE:NusG-CTD interface are highly conserved among different bacteria (Fig. S9), underscoring the importance of this interaction for bacterial viability. Although F165 was proposed to be required for an intramolecular NTD-CTD domain interaction in

Aquifex aeolicus NusG (23), the NusE:NusG-CTD complex structure clearly reveals that F165D or F165T mutations disturb the NusE:NusG-CTD interface.

Coregulation of transcription and translation was initially identified within the attenuation system that controls expression of amino acid biosynthetic operons (24), and polarity was shown to be the result of premature Rho-dependent transcription termination induced by a translation terminating mutation (25). The reported correlation between the rates of transcription and translation of the *infB* and *lacZ* genes, however, implies a direct linkage between the two processes (11). The interaction surface of NusE with NusG ($\sim 1100 \text{ \AA}^2$) is still accessible when NusE forms part of the 30S ribosomal subunit (Fig 3A). Thus, NusE could mediate simultaneous formation of a NusG-NTD:RNAP complex and a NusG-CTD:ribosome complex (Fig 3B). This analysis suggests direct physical coupling of transcription and translation *via* NusG (Fig 3C), a notion that is supported by the observation that NusG depletion decreases translation elongation rates (12).

We next investigated whether the physical coupling of transcription and translation *via* NusG interferes with the known Rho-related functions of NusG (8, 9). HSQC-titrations of ^{15}N -NusG-NTD or ^{15}N -NusG-CTD with Rho showed a nearly complete loss of signals for NusG-CTD, but no effect for NusG-NTD (Fig. S10). The resonances of NusG-CTD are broadened beyond detection by the dramatic increase of the rotational correlation time on NusG-CTD:Rho complex formation (molecular mass of NusG-CTD: 6.9 kDa; molecular mass upon addition of Rho-hexamers: 288.9 kDa). Thus, Rho binds to NusG-CTD, but not to NusG-NTD. Signals of residues from the highly flexible N-terminus of NusG-CTD (R123, K125, T126; Fig. 4A) could be observed in the presence of Rho, suggesting that they are located at the surface of the complex and do not contribute to binding. These residues are part of the linker region between NusG-CTD and NusG-NTD in full-length NusG, and their high flexibility combined with their absence from the NusG-NTD:Rho binding interface indicates that NusG-NTD remains flexibly linked to NusG-CTD in the complex with Rho. Indeed, all detectable signals of full-length ^{15}N -NusG after addition of Rho could be assigned exclusively to residues of NusG-NTD and the linker region. HSQC displacement experiments show the NusG-CTD:Rho and the NusG-CTD:NusE interactions to be mutually exclusive (Fig. 4B, 4C).

Chromatine immunoprecipitation (ChIP) chip analysis indicates that Rho collocates with elongating RNAP soon after transcription initiation (26, 27). How Rho is

recruited to the TEC is not clear. Nevertheless, termination does not occur until transcription reaches the operon terminus. Our findings suggest a straightforward explanation for this delay in termination. We suggest that during coupled transcription-translation NusG-CTD is bound to ribosomal NusE and is therefore unavailable for binding to Rho. Release of ribosomes at the ends of operons frees the NusG-CTD to interact with and stimulate Rho. This mechanism may complement the occlusion of RNA to Rho by translating ribosomes.

The different relative affinities of NusG for NusB:NusE-CTD ($K_d=50 \mu\text{M}$) and for Rho, ($K_d=12 \text{ nM}$; (28)) as reflected in the displacement experiments (Fig. 4) appear to argue against the above model. However, NusE:NusG-CTD interaction takes place within a complex with TEC that includes many other factors. Other interactions with the TEC may significantly lower the NusE:NusG-CTD K_d . For example, during processive antitermination NusB:NusE binds *boxA* RNA (14-16), linking these factors with RNAP and enhancing the overall stability of the antitermination complex. Consistent with the idea that the NusE:NusG interaction is stabilized on RNAP, we found that overproduction of NusG-CTD in wild-type *E. coli* is not toxic, suggesting that it does not efficiently compete with wild type NusG for binding to TEC; (1)). In contrast, NusG-CTD did compete with a NusG mutant with reduced affinity for NusE. Thus, overexpression of NusG-CTD was lethal in a *nusGF165A* mutant strain, in which a key NusE interface residue on NusG is altered. We suggest that NusG-CTD titrates isolated NusE and/or ribosome-bound NusE in the mutant cells.

Finally, our data also explain the puzzling observation that ribosome-bound NusE still supports antitermination (29). According to our model, this activity entails RNAP-ribosome coupling through NusG.

References:

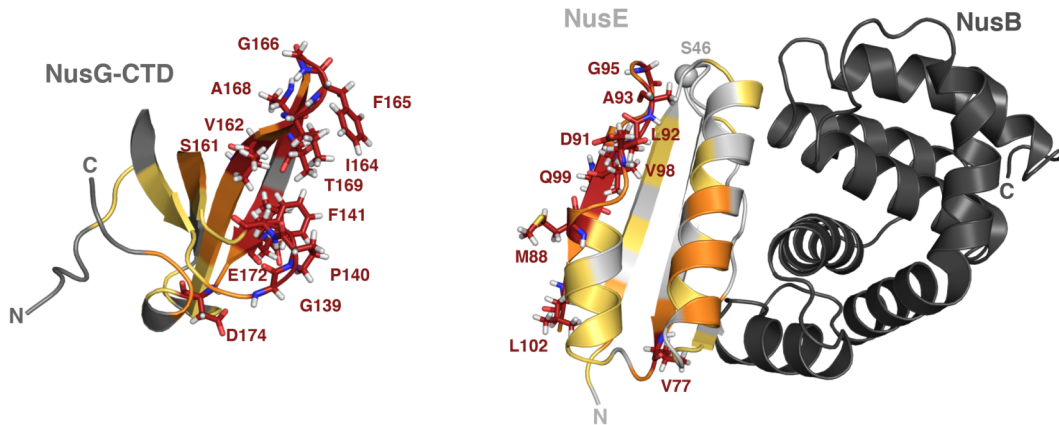
1. R. A. Mooney, K. Schweimer, P. Rösch, M. E. Gottesman, R. Landick, *J. Mol. Biol.* **391**, 341 (2009).
2. L. Downing, S. L. Sullivan, M. E. Gottesman, P. P. Dennis, *J. Bacteriol.* **172**, 1621 (1990).
3. N. C. Kyrpides, C. R. Woese, C. A. Ouzounis, *Trends Biochem. Sci.* **21**, 425 (1996).
4. N. C. Kyrpides, C. A. Ouzounis, *Proc. Natl. Acad. Sci. U. S. A.* **96**, 8545 (1999).
5. T. Steiner, J. T. Kaiser, S. Marinkovic, R. Huber, M. C. Wahl, *EMBO J.* **21**, 4641 (2002).
6. G. A. Hartzog, T. Wada, H. Handa, F. Winston, *Genes Dev.* **12**, 357 (1998).
7. S. W. Mason, J. Greenblatt, *Genes Dev.* **5**, 1504 (1991).
8. S. L. Sullivan, M. E. Gottesman, *Cell* **68**, 989 (1992).
9. C. J. Cardinale *et al.*, *Science* **320**, 935 (2008).
10. G. A. Belogurov, R. A. Mooney, V. Svetlov, R. Landick, I. Artsimovitch, *EMBO J.* **28**, 112 (2009).
11. U. Vogel, K. F. Jensen, *J. Biol. Chem.* **269**, 16236 (1994).
12. M. Zellars, C. L. Squires, *Mol Microbiol* **32**, 1296 (1999).
13. C. L. Squires, D. Zaporojets, *Annu. Rev. Microbiol.* **54**, 775 (2000).
14. S. J. Greive, A. F. Lins, P. H. von Hippel, *J. Biol. Chem.* **280**, 36397 (2005).
15. X. Luo *et al.*, *Mol. Cell* **32**, 791 (2008).
16. B. M. Burmann, X. Luo, M. C. Wahl, P. Rösch, M. E. Gottesman, *Nucleic Acids Res.* **38**, 314 (2010).
17. M. Torres, C. Condon, J. M. Balada, C. Squires, C. L. Squires, *Embo J* **20**, 3811 (2001).
18. J. Greenblatt, J. R. Nodwell, S. W. Mason, *Nature* **364**, 401 (1993).
19. S. L. Sullivan, D. F. Ward, M. E. Gottesman, *J Bacteriol* **174**, 1339 (1992).
20. See the supporting material available on *Science* Online.

21. C. Zwahlen *et al.*, *J. Am. Chem. Soc.* **119**, 6711 (1997).
22. Single-letter abbreviations for amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.
23. J. R. Knowlton *et al.*, *Biochemistry* **42**, 2275 (2003).
24. E. B. Keller, J. M. Calvo, *Proc. Nat. Acad. Sci. USA* **76**, 6186 (1979).
25. S. Adhya, M. Gottesman, *Annu. Rev. Biochem.* **47**, 967 (1978).
26. R. A. Mooney *et al.*, *Mol Cell* **33**, 97 (2009).
27. J. M. Peters *et al.*, *Proc. Natl. Acad. Soc. U S A* **106**, 15406 (2009).
28. Z. Pasman, P. H. von Hippel, *Biochemistry (N. Y.)* **39**, 5573 (2000).
29. A. Das, B. Ghosh, S. Barik, K. Wolska, *Proc Natl Acad Sci U S A* **82**, 4070 (1985).
30. T. M. Schmeing, V. Ramakrishnan, *Nature* **461**, 1234 (2009).
31. We thank Ramona Heissmann for excellent technical assistance. Financial support by the Deutsche Forschungsgemeinschaft to PR (Ro617/16-1) and MCW (Wa1126/3-1), BLS (NSF MCB-0744422) as well as by the NIH to MEG (GM037219) is gratefully acknowledged.

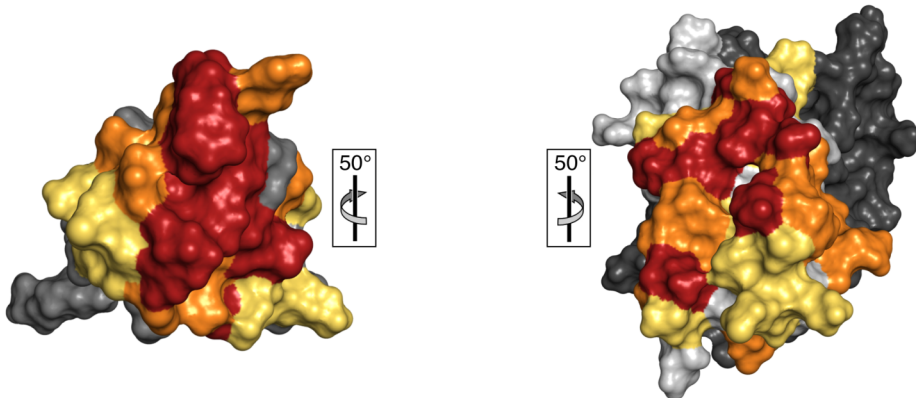
Figures:

Fig. 1:

A



B



C

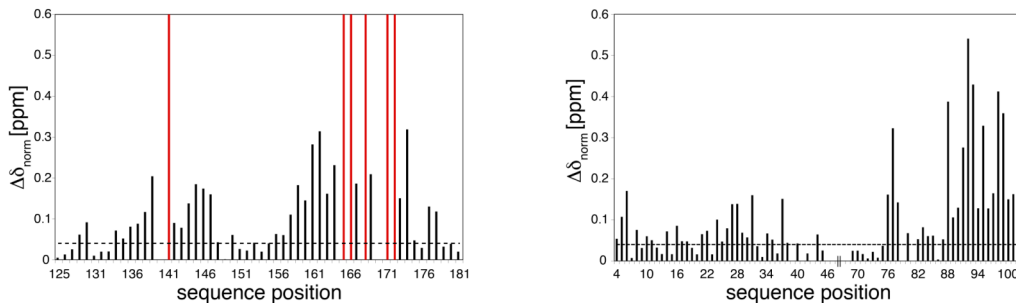
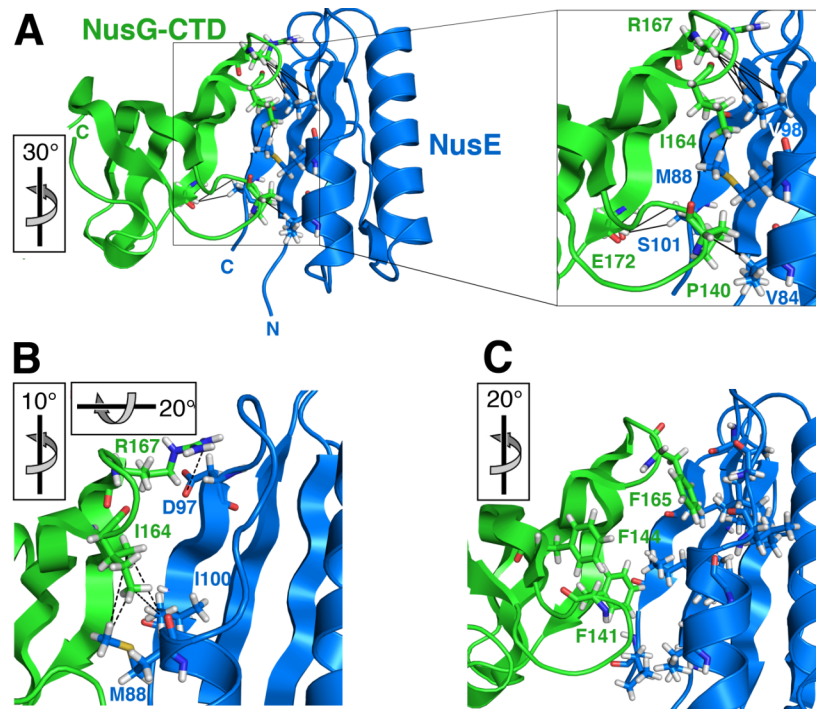


Fig. 1:

A) Mapping of chemical shift changes ($\Delta\delta_{\text{norm}}[\text{ppm}] > 0.2$, red; > 0.1 , orange; > 0.04 , yellow) on the structures of NusG-CTD (gray; Protein Data Bank identification code (PDB ID) 2JVV; (1)) and the NusB:NusE complex (dark gray and light gray, respectively; PDB ID 3D3B; (15)). Strongly affected residues are labeled and shown as sticks. Coloring scheme: carbon, red; hydrogen, white; oxygen, red; nitrogen, blue; sulphur, yellow. The light gray sphere in NusE denotes the Ca position of S46, which in this construct

replaces residues 46 to 67 of wild-type NusE (20). Numbering is based on the residue numbers of wild-type NusE (1 to 46 and 68 to 103).

- B) Surface representation of the structures shown in A). Orientations relative to A) are indicated.
- C) HSQC-derived chemical shift changes vs. sequence position; left: NusG-CTD chemical shift changes on titration with NusB:NusE; right: NusE (in the NusB:NusE complex) chemical shift changes on titration with NusG-CTD; missing residues of the NusE ribosome-binding loop are indicated by a small double-bars on the sequence axis. Dotted line: significance level of $\Delta\delta_{\text{nom}}[\text{ppm}] = 0.04$; red bars: NusG-CTD signals disappearing on complex formation.

Fig. 2:**Fig. 2:**

- A) Experimental basis for the structure determination of the NusE:NusG-CTD complex. For clarity, NusB was omitted. Coloring scheme: carbon, marine (NusE) and green (NusG); hydrogen, white; oxygen, red; nitrogen, blue; sulphur, yellow. Black lines represent unambiguously assigned NOEs between NusG-R167-H δ :NusE-V98-H γ , NusG-I164-H δ :NusE-M88-H ϵ , NusG-E172-H β :NusE-S101-H β , and NusG-P140-H α :NusE-V84-H β . These intermolecular NOEs allowed unambiguous determination of the relative orientation of the two proteins by rigid body minimization.
- B) Possible interactions in the NusE:NusG-CTD interface between negatively charged NusE-D97 and positively charged NusG-R167. NusG-I164 is sandwiched by the hydrophobic sidechains of NusE-M88 and NusE-I100.
- C) NusG-F165 and NusG-F141 bind to a NusE hydrophobic pocket. NusG-F144 is in close proximity to the interaction site and possibly participates indirectly in this interaction. Several NusE hydrophobic sidechains are close to NusG phenylalanine residues. The orientation relative to Fig. 1A is indicated (PDB ID: 2KVQ).

Fig. 3:

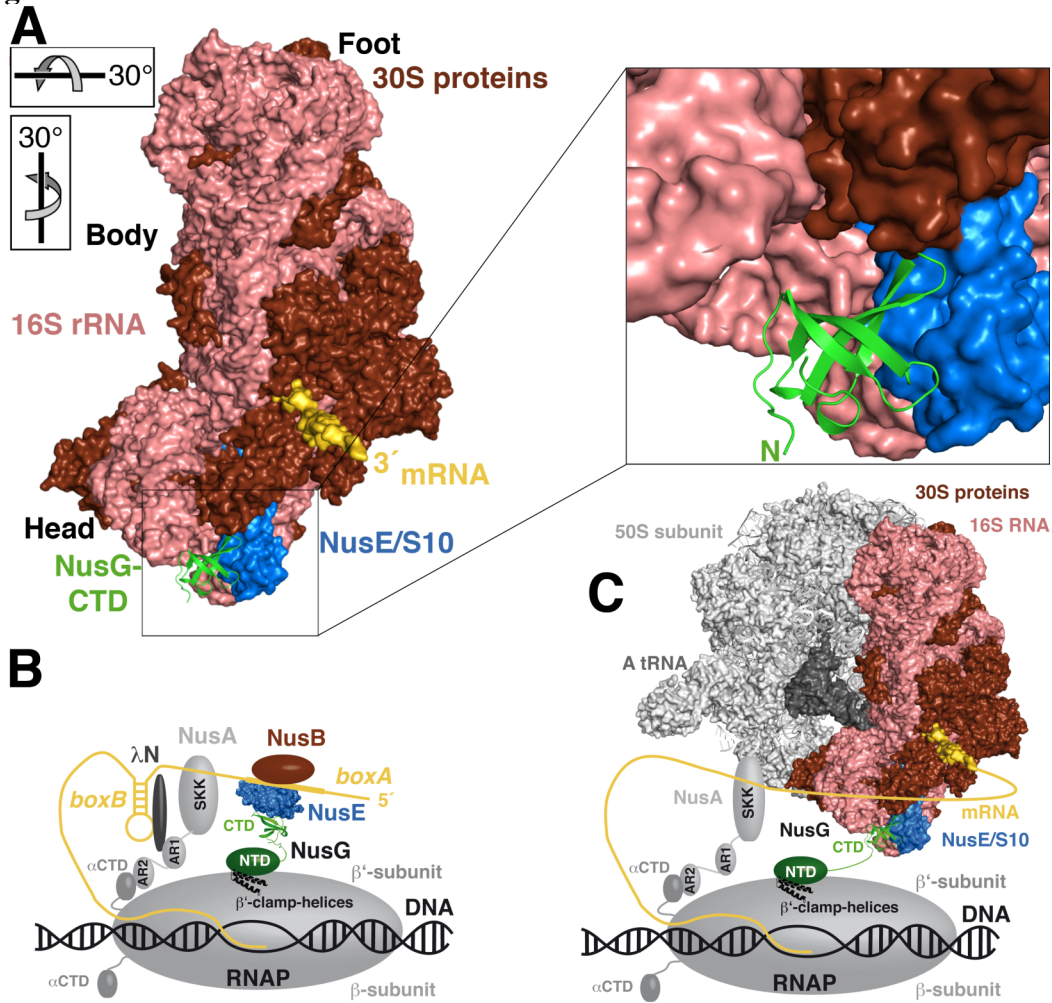


Fig. 3:

- A. Possible interaction between NusG and NusE/S10 in the 30S ribosomal subunit. The NusE (marine) NusG-CTD (green) heterodimer was aligned to the structure of the *E. coli* 30S ribosomal subunit (ribosomal proteins, brown; 16S rRNA, rose). Landmark features of the 30S subunit (head, body, foot) are labeled. S10 is part of the head region of the 30S subunit in close proximity to the entrance site of the mRNA. The path of the mRNA is shown in gold. The orientation relative to Fig. 1A is shown.
- B. Schematic representation of the assembly of the λ N mediated antitermination complex (20). NusA domains AR1 (acidic repeat 1), AR2 (acidic repeat 2), and SKK (S1-KH1-KH2) are in light gray; RNAP α CTD, gray; *boxA*, *boxB*, and spacer region of *nut* RNA, gold; λ N, dark gray; NusB, brown; NusE, marine; NusG-CTD, green; NusG-NTD, dark green; β '-clamp-helices of the RNAP, black.

- C. Model of the physical coupling between transcription and translation. Coloring as in A) and B), the 50S ribosomal subunit is shown in light gray and the A-site tRNA in dark gray. NusG links RNAP and the ribosome so that efficient and fast translation of the nascent mRNA can occur. Coordinates for the complete 70S ribosome, the mRNA and the A-site tRNA were kindly provided by T.M. Schmeing and V. Ramakrishnan, MRC, Cambridge, UK (30).

Fig. 4:

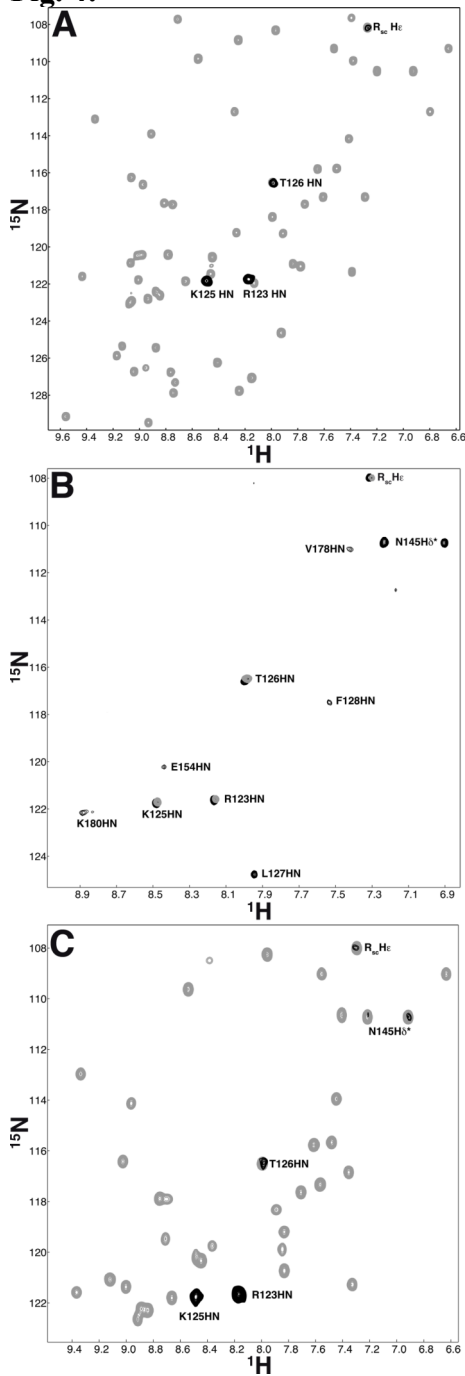


Fig. 4:

^1H , ^{15}N HSQC-spectra of ^{15}N -NusG-CTD titrated with Rho. Spectrum of free NusG-CTD, gray, and with an additional equivalent of Rho, black. All gray signals disappeared after Rho addition, and the remaining signals (black) are part of the spacer region between the NusG domains.

^1H , ^{15}N HSQC-spectra of a displacement titration of a 1:1 of ^{15}N -NusG-CTD:Rho complex (gray) with NusB:NusE. Addition of 20 equivalents of NusB:NusE, black. All newly appearing signals belong either to the spacer and the C-

terminus of NusG or are on the side of the spacer region pointing away from the interaction surface (E154, N145).

^1H , ^{15}N HSQC-spectra of a displacement titration of a 1:1 ^{15}N -NusG-CTD:NusB:NusE complex (gray) with Rho. Addition of one equivalent of Rho, black. The resulting spectrum is identical to the spectrum observed with the NusG-CTD:Rho complex (part (A) of this figure).

Supporting Online Material:**Experimental Procedures:**

NusE^{Δloop}:NusB-complex: The gene encoding NusE^{Δloop} was cloned *via* BamH1 and EcoR1 restriction sites into the *E. coli* expression vector pGEX-6P (GE Healthcare, Munich, Germany) (1). The gene for NusB was cloned *via* BamH1 and NdeI restriction sites into the *E. coli* expression vector pET29b (Novagen, Madison, WI, USA) (2). Expression and purification of the dimeric complex was as described (2). The identity and structural integrity of the purified protein complex was analyzed by SDS-PAGE and NMR spectroscopy.

NusG-CTD: NusG-CTD (123-181) was cloned *via* BamH1 and NcoI into the *E. coli* expression vector pETGB1A (G. Stier, EMBL, Heidelberg, Germany) containing an aminoterminal GB1-fusion tag (streptococcal immunoglobulin-binding domain of protein G), an aminoterminal hexahistidine tag, and a TEV cleavage site between GB1 and NusG-CTD. *E. coli* strain BL21(DE3) (Novagen, Madison, WI, USA) harboring the recombinant plasmid was grown at 37 °C in LB medium containing kanamycin (30 µg/ml) to OD₆₀₀ = 0.8 and then induced with isopropyl 1-thio-β-D-galactopyranoside (IPTG; final concentration 1 mM). Cells were harvested 4 h after induction, resuspended in 4 times the cell weight of lysis buffer (50 mM TRIS, 150 mM NaCl, pH 7.5, 1 protease inhibitor tablet (Complete, EDTA-free, (Roche Diagnostics GmbH, Mannheim, Germany)), and lysed with a micro-fluidizer (Microfluidics, Newton, MA, USA). The lysate was centrifuged for 30 min at 12,000 g, and was applied to a Ni²⁺-NTA HiTrap column (GE Healthcare, Munich, Germany) and eluted *via* a step gradient with increasing imidazole concentrations (10 mM – 500 mM). The eluted fractions containing the GB1-NusG-CTD fusion protein were cleaved by TEV-protease during dialysis against 50 mM TRIS, 150 mM NaCl, pH 7.5 at 4 °C overnight. The cleaved protein was reapplied to a Ni²⁺-NTA HiTrap column. The flow-through (containing NusG-CTD) was concentrated by ultrafiltration (Vivascience, MWCO 5,000 Da) and applied to a Superdex-75 gel filtration column (50 mM TRIS-HCl, 150 mM NaCl, pH 7.5; GE Healthcare, Munich, Germany). The fractions containing NusG-CTD were pooled and dialyzed against 25 mM HEPES, 50 mM NaCl, pH 7.5. The identity and structural integrity of the purified protein was analyzed by SDS-PAGE and NMR spectroscopy.

Rho: *E. coli* Rho was expressed and purified as described (3).

Isotope labeled proteins: ¹⁵N- and ¹⁵N-, ¹³C-labeled proteins were obtained by growing *E. coli* in M9-minimal media (4) upon respective addition of (¹⁵NH₄)₂SO₄ (Campro Scientific, Berlin,

Germany) and ^{13}C D-glucose (Spectra Stable Isotopes, Columbia, MD, USA; (5)). For the production of ^2H - ^{15}N - ^{13}C -NusE, four precultures with increasing concentrations of D_2O (Euriso-top, Gif-sur-Yvette, France) were used. Expression and purification was as described before.

Analytical size exclusion chromatography: For analytical size exclusion chromatography, the experimental procedure was adapted from Trowitzsch *et al.* (6). NusB:NusE and NusG proteins were mixed in approximately equimolar ratios, applied on a Superdex-75 PC 3.2 column (GE Healthcare, Munich, Germany), and chromatographed in 10 mM TRIS-HCl, 50 mM NaCl, 2 mM DTT, pH 7.5 using a SMART protein purification system (GE Healthcare, Munich, Germany). For a typical run, 30 μl of sample were loaded on the column at a flow rate of 40 $\mu\text{l}/\text{min}$. 40 μl fractions were collected and analyzed by 15 % SDS-PAGE.

NMR: All NMR experiments were performed in NMR-buffer containing 25 mM HEPES, 50 mM NaCl, pH 7.5 (for NusG:NusE:NusB interactions) and 10 mM TRIS-HCl, 50 mM KCl, 1 mM MgCl_2 , 0.1 mM EDTA, pH 7.7 (for NusG:Rho interactions). The experiments were recorded at 298 K for the titrations and 310 K for assignment experiments on Bruker Avance 700 MHz and Avance 800 MHz spectrometers with cryogenically cooled triple-resonance probes equipped with pulsed field-gradient capabilities.

For resonance assignment of NusG-CTD, standard double and triple resonance through-bond experiments were recorded (7) and for NusE the following TROSY (transverse relaxation optimized spectroscopy)-type NMR experiments (8, 9) were additionally performed: [^{15}N , ^1H]-TROSY, tr-HN(CA)CB, tr-HNCA. For obtaining distance restraints, ^{15}N -separated and ^{13}C -separated nuclear Overhauser enhancement (NOE) experiments were recorded (10, 11). Intermolecular NOEs were recorded with one partner $^{15}\text{N}/^{13}\text{C}$ -labeled and the other unlabeled in D_2O using three-dimensional ^{13}C -separated/ ^{12}C -filtered NOE experiments (12). Distance restraints for structure calculation were semiquantitatively derived from cross peak intensities of 3D ^{13}C -edited NOESY experiments (mixing time 100 ms) and 3D ^{13}C -filtered (F1), ^{13}C -edited (F2) NOESY experiments (mixing time 120 ms). In addition 3D ^{15}N -edited NOESY experiments (mixing time 60-80 ms) were recorded. These spectra were compared with ^{13}C -edited and ^{15}N -edited NOESY data of the free proteins to validate the assumption that the overall fold of the individual domains does not change upon complex formation. A 3D ^{13}C -filtered (F1), ^{13}C -edited (F2) NOESY experiment was also recorded for the isolated NusE:NusB complex to identify intermolecular NOEs between NusE and NusB before analyzing data of the ternary complex.

The observed intermolecular NOEs were grouped in two distance classes with upper distance limits of 4.0 Å (strong) or 5.0 Å (weak). NOEs involving diastereotopic methyl groups (Val, Leu) were represented by a $(\sum r^{-6})^{-1/6}$ sum over all six protons to avoid bias by spin diffusion. Assignment of NOE cross signals were facilitated by the amino acid composition in the complex interface, because NusG-CTD contributes three phenylalanines whereas NusE does not contain any aromatic residue close to the interface. Intermolecular NOEs involving an aromatic residue (Phe141, Phe144 or Phe165) were always used as ambiguous restraints with the representation by a $(\sum r^{-6})^{-1/6}$ sum over all protons and all three residues to avoid bias by spin diffusion or wrong assignment to individual protons of the aromatic side chains.

NMR data were processed using in-house routines and visualized with NMRView (13). Dissociation constants were determined from chemical shift changes of residues showing fast exchange behavior using a two-state model.

Modeling of the complex: The structural model was generated using rigid body docking with flexible side chains of residues close to the binding region (14). In the calculation, we only used the intermolecular distance restraints based on NOESY experiments (Table S1) as experimental restraints. Docking was carried out using XPLOR-NIH 2.12 (15) with a modified version of the protocol for docking of protein complexes (dock_tor_rigid.py taken from XPLOR-NIH supplied protocols; (14)). Changes of the original protocol include neglect of residual dipolar couplings and a potential term for the radius of gyration. Starting structures were taken from the high resolution crystal structure of the *E. coli* NusB:NusE complex (PDB: 3D3B; (1)) and the solution structure of *E. coli* NusG-CTD (PDB: 2JVV; (16)). Because NusB did not interact with NusG-CTD in the NusB:NusE complex, coordinates of NusB were omitted during calculation. In addition, the seven amino-terminal residues of NusE, which showed no regular secondary structure and exhibited a high degree of flexibility in the crystal (1), were deleted *in silico* to avoid steric clashes. During initial calculations all side chains were permitted to rotate freely. For the final calculation, full side chain flexibility was allowed for those residues that showed an average distance below 8 Å for at least one atom to any atom of the partner molecule in earlier rounds of calculations. 120 structures were calculated in the final iteration. Without exception, all structures exhibit the same orientation of the proteins (backbone atomic rmsd 0.49 Å over all residues, no distance restraint violation > 0.5 Å, Table S1) demonstrating that the intermolecular distance restraints are sufficient for structure determination with rigid body minimization.

Ribosome Binding of NusE and NusE^{Δloop}: Binding of GST-NusE and GST- NusE^{Δloop} to ribosomes was tested as described previously for other ribosomal proteins and their variants (17). DNA

fragments encoding NusE or NusE^{Δloop} were cloned into the BamHI and EcoRI restriction sites of pGEX-6P-1 (GE Healthcare) for expression as N-terminal GST fusion proteins (1). *E. coli* BL21(DE3) cells were transformed with the plasmids and grown at 32 °C in 200 ml LB medium with 100 μg/ml ampicillin. At OD₄₅₀ = 0.1, IPTG was added to 0.5 mM final concentration. At OD₄₅₀ = 1.5, cells were harvested by centrifugation. The cell pellets were washed once with 1 ml buffer A (20 mM HEPES-KOH, 6 mM MgCl₂, 30 mM NH₄Cl, 6 mM β-mercaptoethanol, pH 7.5), resuspended in 2 ml of buffer A and split into two 1-ml aliquots. After addition of lysozyme (100 μl of a 15 mg/ml solution) and incubation on ice for 3 min, cell lysis was completed by freeze-thawing. Lysates were clarified by spinning at 23,000 rpm for 30 min in an S100-AT4 rotor (Thermo Scientific). The duplicate supernatants from each culture were pooled and centrifuged for 4 hours at 43,000 rpm in an S100-AT4 rotor. The pellets were resuspended overnight in 200 μl of buffer A and centrifuged at 7900 x g in a Fresco 17 centrifuge (Heraeus) for 10 min at 4 °C. The resulting supernatants contain 'crude' ribosomes. 150 μl of the crude ribosome preparations were mixed with 1.8 ml buffer B (20 mM HEPES-KOH, 30 mM MgCl₂, 1 M NH₄Cl, 6 mM β-mercaptoethanol, pH 7.5), incubated for 1 hour at 4 °C and then centrifuged for 4 hours at 53,000 rpm in an S100-AT4 rotor. The pellets were rinsed once with 200 μl buffer A, resuspended overnight in 100 μl buffer A and centrifuged at 7900 x g in a Fresco 17 centrifuge for 10 min at 4 °C. The resulting supernatants contain 'salt-washed' ribosomes.

Proteins from ribosomes (0.1 A₂₆₀ equivalents) were separated by 12 % SDS PAGE and electroblotted on a nitrocellulose transfer membrane (Whatman). For probing GST-tagged NusE or NusE^{Δloop}, the membrane was first incubated with a rabbit anti-GST antibody (Invitrogen) and subsequently by a goat anti-rabbit IgG (Dianova). The signal was detected by using ECL Western blotting detection reagents (GE Healthcare) and exposing a high performance chemiluminescence film (GE Healthcare). The film was developed on a KONICA developer.

Assembly of the λ N mediated antitermination complex

The highly conserved bacterial transcription elongation factor NusA (18) consists of the following functional subdomains: an N-terminal domain that interacts with RNAP (19), three RNA-binding domains, S1, KH1 and KH2 (20-22) forming the SKK-domain (23), and two C-terminal acidic domains AR1 and AR2 (24-26). NusA participates in the initiation-to-elongation transition due to its interaction with the α CTD of the RNAP mediated by NusA-AR2 (19). Furthermore, its SKK-domain interacts with the spacer region of the *nut* RNA (23), and the N-terminal arginine-rich-motif of λ N binds to the *boxB* of the *nut* RNA (27, 28), where it induces helix formation in λ N (29, 30). Additionally, λ N recruits NusA-AR1 to this site (24, 26). NusB loads NusE onto *boxA* of *nut* RNA (1, 2, 31), where the NusB:NusE heterodimer forms an extended mosaic RNA-binding interface (1). NusG-CTD interaction with the NusB:NusE heterodimer leads to the formation of a closed and compact complex due to the simultaneous interaction of the NusG-NTD with the β' -clamp-helices of the RNAP (32). These β' -clamp-helices are an important regulatory site on the RNAP as in addition to providing binding sites for NusG and its paralog RfaH (32, 33) they form an interaction site with σ -factor on RNAP (34, 35). Whether NusG or RfaH is able to displace σ -factor from this site remains unknown as so far only models for the initiation-to-elongation transition exist (36).

Ribosome Binding of NusE and NusE Δ^{loop}

In NusE Δ^{loop} (ribosomal protein S10 Δ^{loop}), residues 46-67 are replaced by a serine residue (1). Because residues 46-67 of NusE form a loop, which in the ribosome-bound form of the protein penetrates the 30S ribosomal subunit (37), we reasoned that NusE Δ^{loop} may fail to bind to ribosomes and therefore fail to support translation. In order to test this idea, we directly monitored binding of GST-NusE and GST-NusE Δ^{loop} to ribosomes. Since the N-terminus of NusE is accessible on the surface of the 30S ribosomal subunit (37), an N-terminal GST fusion should not interfere with stable ribosome incorporation of the protein. Indeed, GST-NusE was incorporated readily and in a salt-stable manner into ribosomes (Figure S2, lanes 1-4). In contrast, while residual amounts of GST-NusE Δ^{loop} were seen associated with crude ribosome preparations, the truncated fusion protein was completely lacking from salt-washed ribosomes (Figure S2, lane 5-8). Thus, NusE behaves differently from some other ribosomal proteins, such as L4, in which analogous ribosome-penetrating loops are not required for stable ribosome association (17).

Table S1: Summary of structure calculation. Restraint violation and coordinate precision are calculated for all 120 structures. Values in parentheses are for the best 40 structures.

Experimental restraints	
Intermolecular distance restraints	17
Restraint violations	
Maximal violation	0.45 (0.14) Å
Average violation	0.017 (0.012) Å
Coordinate precision	
Backbone	0.48 (0.39) Å

Supporting Figures:

Fig. S1:

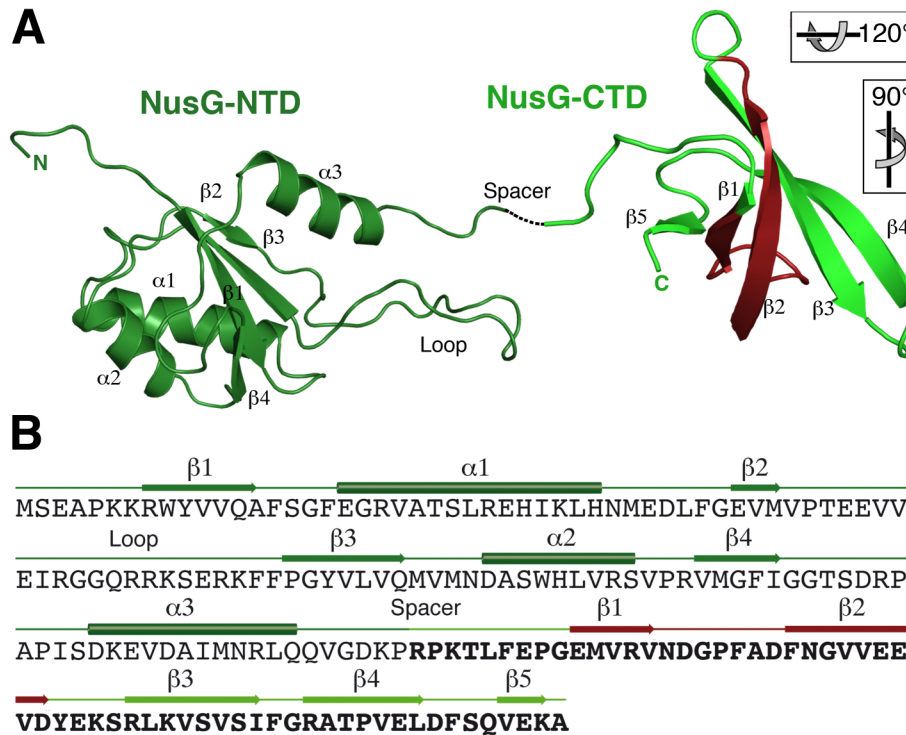
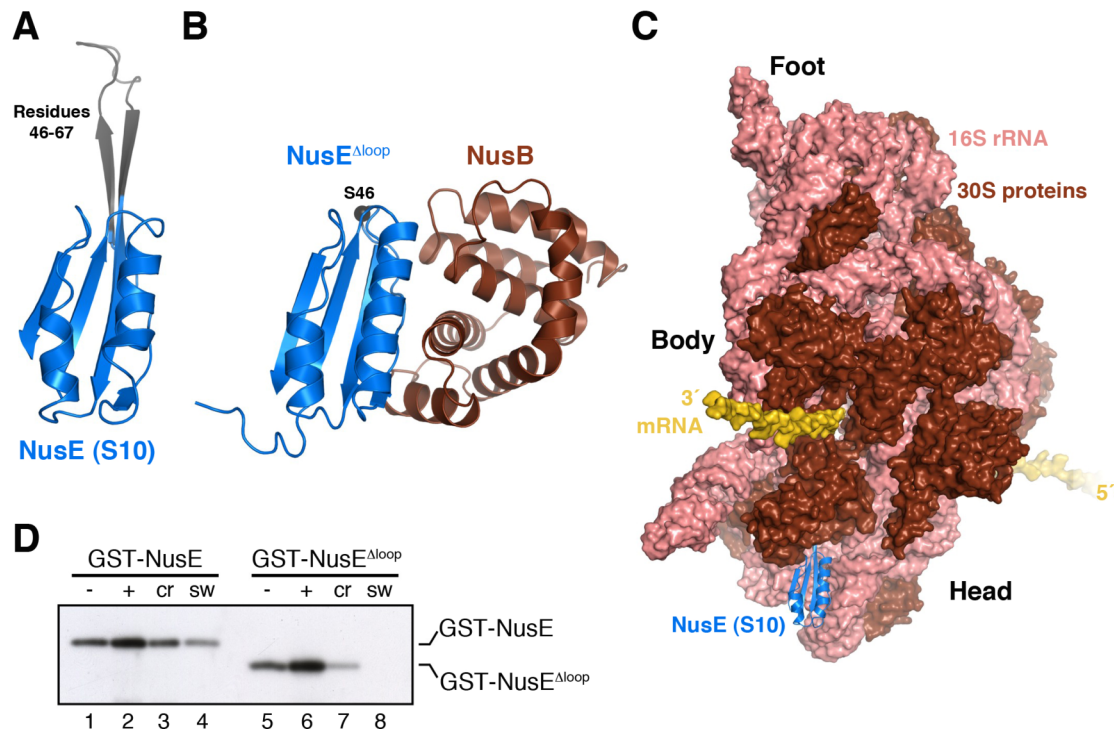


Fig. S1:

- A. The structure of *E. coli* NusG. NusG-NTD (1-116; dark green; PDB ID 2K06; (16)) is required for RNAP interaction (32). NusG-CTD (132-181; green; PDB ID: 2JVV; (16)) is connected with the NTD via a flexible linker region (117-131) and contains a KOW-motif (red; $\beta 1$, $\beta 2$ and connecting loop), which may be involved in protein and/or nucleic acid interactions (38, 39). Termini and secondary structure elements are labeled, and the orientation relative to Fig. 1A is shown.
- B. Amino-acid sequence of full-length NusG with corresponding secondary structure elements as shown in A). Key amino-acids of the NusG-CTD (123-181) construct used in this study are shown in bold.

Fig. S2:

Fig. S2: Ribosome Binding of NusE and NusE^{Δloop}

- A. The structure of ribosome-bound NusE (ribosomal protein S10; from PDB ID 2AVY; (40)). The black region denotes the long, ribosome-binding loop (residues 46-67), which was deleted in NusE^{Δloop}. In this and the following panels NusE or NusE^{Δloop} are oriented as in Figure 1A.
- B. Ribbon plot of the NusE^{Δloop}-NusB complex (PDB ID 3D3B; (1)). The black sphere in NusE^{Δloop} denotes the C α position of S46, which in NusE^{Δloop} replaces residues 46-67 of full-length NusE.
- C. Position of NusE in the 30S subunit. Model coordinates of the 30S subunit with mRNA were kindly provided by T.M. Schmeing and V. Ramakrishnan, MRC, Cambridge, UK (41). 30S subunit protein - brown; 16S rRNA - rose; mRNA - gold; NusE - marine. Landmark features of the 30S subunit (head, nose, body, toe) are labeled. NusE forms part of the head of the 30S subunit with the long loop deeply buried in the particle interior.
- D. Western blot probing the binding of GST-NusE and GST-NusE^{Δloop} to ribosomes. Equal amounts of cells before (-; lanes 1 and 5) and after (+;

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

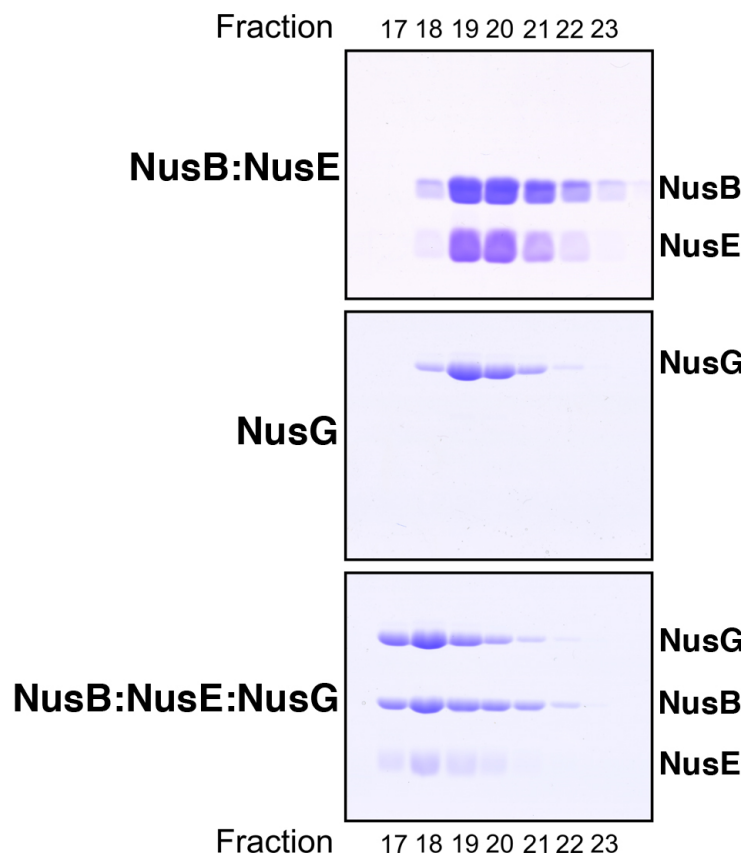
Fig. S3:

Fig. S3: Gel filtration analysis. The panels show the migration of NusB:NusE complex (top), of NusG (middle), and of a mixture of NusB:NusE and NusG (bottom). Fractions collected from each run were analyzed by SDS-PAGE. Equivalent fractions (17-23) are aligned below each other in the three panels. The mixture of NusB:NusE and NusG elutes 1-2 fractions earlier compared to either NusB:NusE or NusG alone, indicating formation of a NusB:NusE:NusG complex.

Fig. S4:

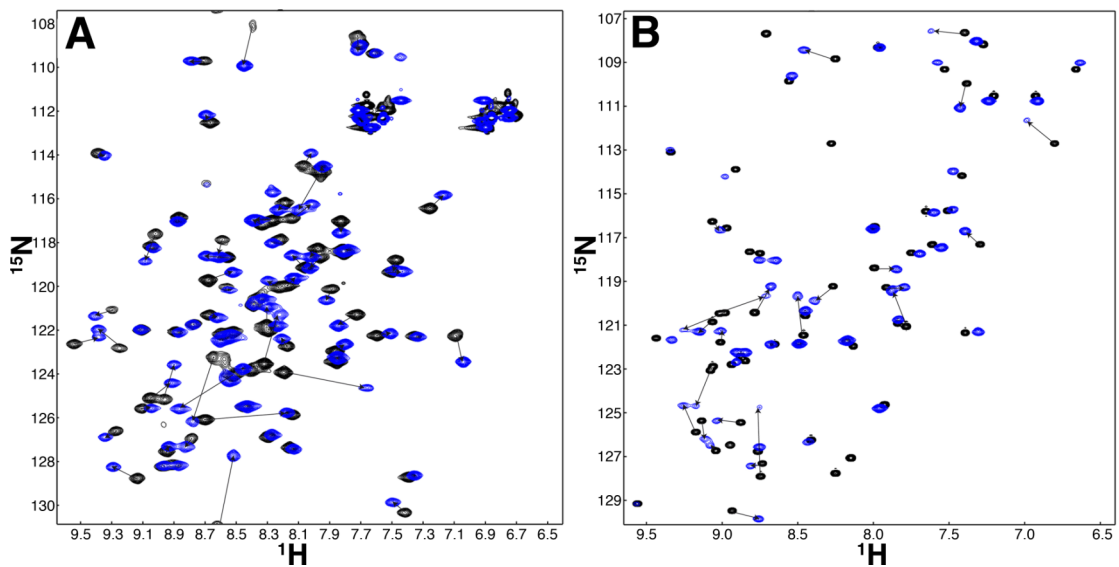
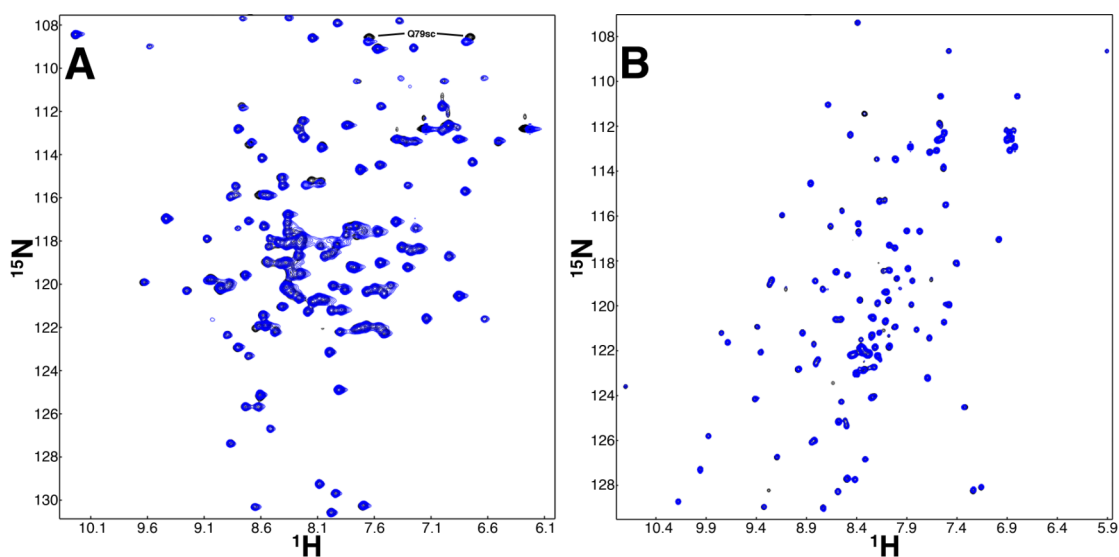


Fig. S4:

- A. ^1H , ^{15}N HSQC-spectra of NusB: ^{15}N -NusE titrated with NusG-CTD. Spectrum of free NusB: ^{15}N -NusE, 420 μM , black; spectrum of NusB: ^{15}N -NusE, 270 μM , in the presence of NusG-CTD, 410 μM , blue. Arrows indicate chemical shift changes.
- B. ^1H , ^{15}N HSQC-spectra of ^{15}N -NusG-CTD titrated with NusB:NusE. Spectrum of free NusG, 500 μM , black; spectrum of ^{15}N -NusG-CTD, 260 μM , in the presence of NusB:NusE, 390 μM , blue. Arrows indicate chemical shift changes.

Fig. S5:**Fig. S5:**

- A. ^1H , ^{15}N HSQC-spectra of ^{15}N -NusB:NusE titrated with NusG-CTD. Spectrum of free ^{15}N -NusB:NusE, 500 μM , black; spectrum of ^{15}N -NusB:NusE, 225 μM , in the presence of NusG-CTD, 450 μM , blue. The few slightly shifting signals belong to the NusE:NusB interface. The sidechain of NusB-Q79 is indicated. It is in close proximity of NusE-D19, which shows a slight chemical shift as well. These observed chemical shifts on the NusB side are only indicative for a minor rearrangement of sidechains and slight differences of the chemical environment of a few backbone amides.
- C. ^1H , ^{15}N HSQC-spectra of ^{15}N -NusG-NTD titrated with NusB:NusE. Spectrum of free ^{15}N -NusG-NTD, 350 μM , black; spectrum of ^{15}N -NusG-NTD, 190 μM , in the presence of NusB:NusE, 370 μM , blue.

Fig. S6:

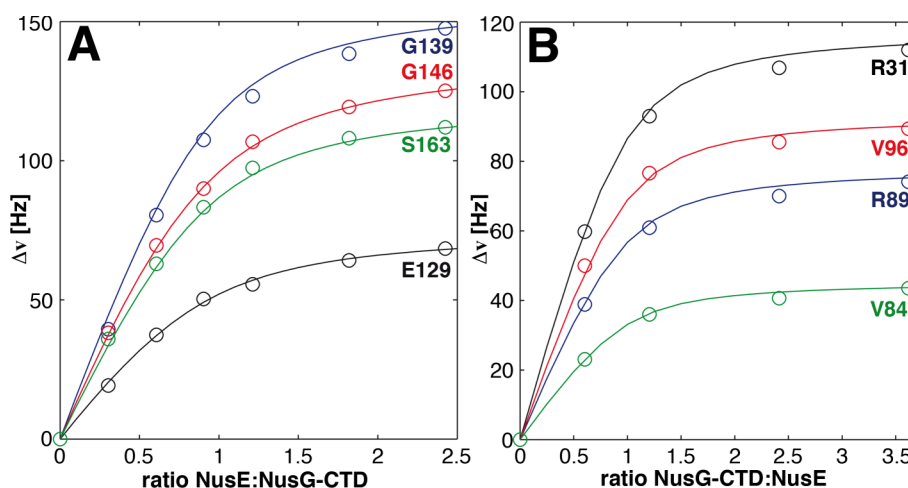


Fig. S6: Backbone amide chemical shift perturbations for selected residues upon titrating unlabeled NusB:NusE to ^{15}N -NusG-CTD (A; starting concentration 500 μM) and NusG-CTD to NusB: ^{15}N -NusE (B; starting concentration 420 μM). The chemical shifts were monitored with ^1H , ^{15}N HSQC spectra. The lines represent nonlinear least squares best fits of the normalized changes in the ^1H and ^{15}N chemical shifts to all of the titration data simultaneously, using a bimolecular equilibrium binding model. The optimized value of the equilibrium dissociation constant is $54 \pm 10 \mu\text{M}$ for A and $40 \pm 2 \mu\text{M}$ for B.

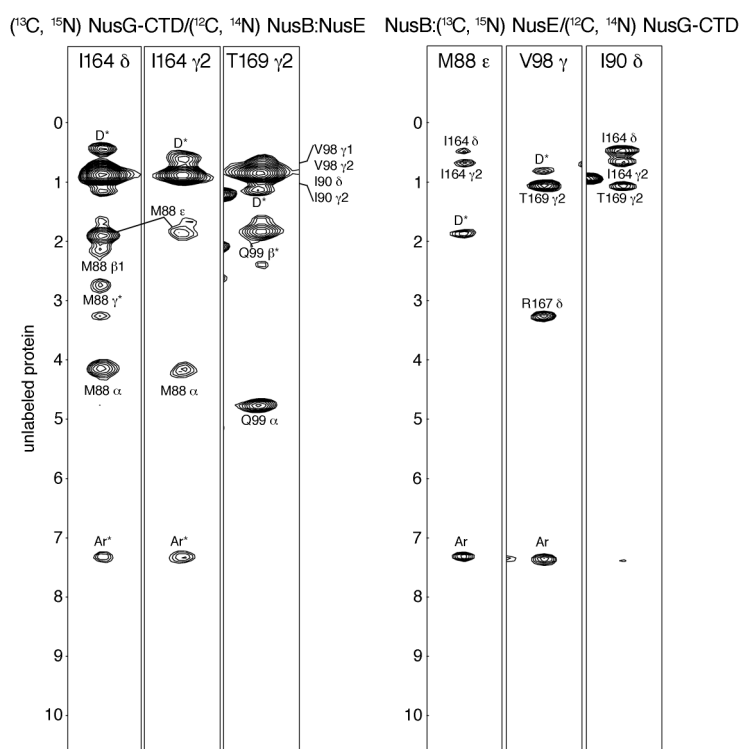
Fig. S7:

Fig. S7: Intermolecular NOEs of the NusB:NusE-NusG CTD complex. Selected strip extracts of 3D ^{13}C -filtered (F1)- ^{13}C -edited (F2) NOE experiments (mixing time 120 ms) of samples containing ^{13}C , ^{15}N labeled NusG-CTD (A) or ^{13}C , ^{15}N labeled NusE (B) in D_2O buffer. Ar indicates only ambiguously defined aromatic peaks. Incompletely suppressed peaks are labeled Ar* (aromatic peaks for NusG-CTD, because NusE is devoid of aromatic residues) and D* (diagonal peaks), respectively.

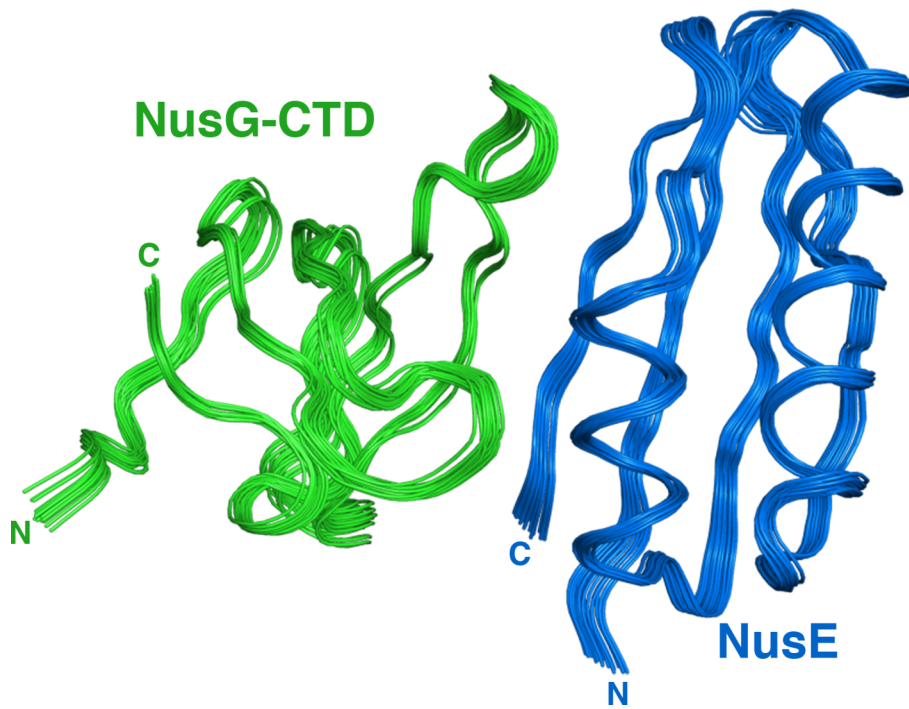
Fig. S8:

Fig. S8: Overlay of 20 calculated structures based on the intermolecular distance restraints as in Table S1. Structures exhibit an rmsd of 0.48 Å (backbone atoms). The orientation is as in Fig. 1A.

Fig. S10:

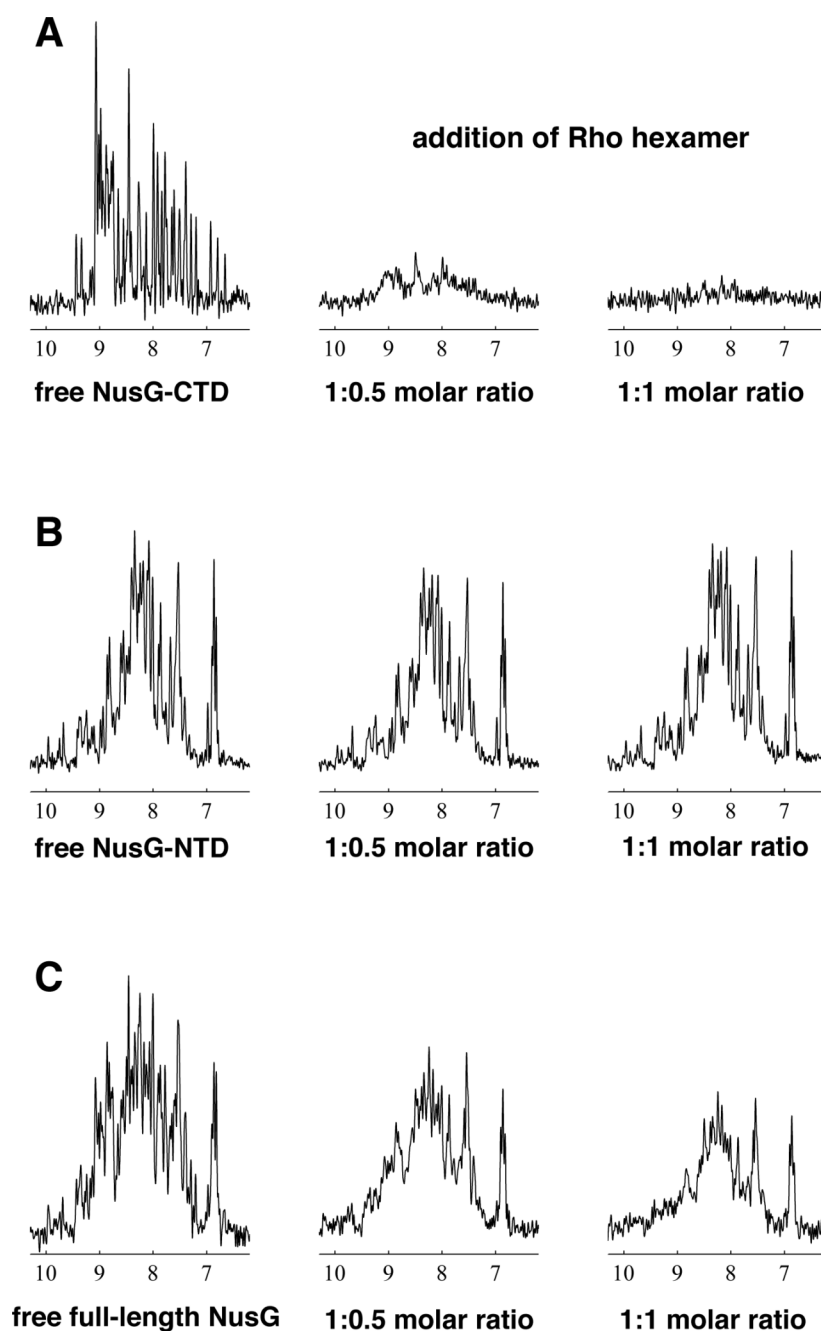


Fig. S10: First 1D section of a ^1H , ^{15}N HSQC spectrum of ^{15}N -NusG-CTD (A), ^{15}N -NusG-NTD (B), titrated with Rho. Left: free NusG-CTD (50 μM); center and right: mixtures containing, respectively, 1:0.5 and 1:1 molar ratios of NusG-CTD:Rho hexamer. Loss of signal for NusG-CTD upon increasing Rho concentration indicates that binding takes place *via* this domain. NusG-NTD shows no significant effect. Full-length ^{15}N -NusG (C) shows significant loss of signal, mainly due to disappearing NusG-CTD signals. The 1D spectra were normalized to standard intensity.

References:

1. X. Luo *et al.*, *Mol. Cell* **32**, 791 (2008).
2. B. M. Burmann, X. Luo, M. C. Wahl, P. Rösch, M. E. Gottesman, *Nucleic Acids Res.* **38**, 314 (2010).
3. R. J. Browne, E. W. Barr, B. L. Stitt, *J. Biol. Chem.* **280**, 13292 (2005).
4. J. Sambrook, E. Fritsch F, T. Maniatis, *Molecular Cloning - A Laboratory Manual* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1994).
5. J. Marley, M. Lu, C. Bracken, *J. Biomol. NMR* **20**, 71 (2001).
6. S. Trowitzsch, G. Weber, R. Lührmann, M. C. Wahl, *J. Biol. Chem.* **283**, 32317 (2008).
7. M. Sattler, J. Schleucher, C. Griesinger, *Prog. Nucl. Magn. Reson. Spectrosc.* **34**, 39 (1999).
8. K. Pervushin, R. Riek, G. Wider, K. Wüthrich, *Proc. Natl. Acad. Sci. U. S. A.* **94**, 12366 (1997).
9. M. Salzmann, K. Pervushin, G. Wider, H. Senn, K. Wüthrich, *Proc. Natl. Acad. Sci. U. S. A.* **95**, 13585 (1998).
10. S. Talluri, G. Wagner, *J. Magn. Reson. B* **112**, 200 (1996).
11. M. Ikura *et al.*, *Biochemistry* **30**, 9216 (1991).
12. C. Zwahlen *et al.*, *J. Am. Chem. Soc.* **119**, 6711 (1997).
13. B. A. Johnson, *Methods Mol. Biol.* **278**, 313 (2004).
14. G. M. Clore, C. D. Schwieters, *J. Am. Chem. Soc.* **125**, 2902 (2003).
15. C. D. Schwieters, J. J. Kuszewski, N. Tjandra, G. M. Clore, *J. Magn. Reson.* **160**, 66 (2003).
16. R. A. Mooney, K. Schweimer, P. Rösch, M. E. Gottesman, R. Landick, *J. Mol. Biol.* **391**, 341 (2009).
17. J. M. Zengel, A. Jerauld, A. Walker, M. C. Wahl, L. Lindahl, *RNA* **9**, 1188 (2003).
18. D. I. Friedman, L. S. Baron, *Virology* **58**, 141 (1974).
19. T. F. Mah, K. Kuznedelov, A. Mushegian, K. Severinov, J. Greenblatt, *Genes Dev.* **14**, 2664 (2000).

20. B. Beuth, S. Pennell, K. B. Arnvig, S. R. Martin, I. A. Taylor, *EMBO J.* **24**, 3576 (2005).
21. M. Worbs, G. P. Bourenkov, H. D. Bartunik, R. Huber, M. C. Wahl, *Mol. Cell* **7**, 1177 (2001).
22. B. Gopal *et al.*, *J. Mol. Biol.* **314**, 1087 (2001).
23. S. Prash *et al.*, *Nucl. Acids Res.* **37**, 4736 (2009).
24. I. Bonin *et al.*, *Proc. Natl. Acad. Sci. U. S. A.* **101**, 13762 (2004).
25. A. Eisenmann, S. Schwarz, S. Prash, K. Schweimer, P. Rösch, *Protein Sci.* **14**, 2018 (2005).
26. S. Prash *et al.*, *Biochemistry* **45**, 4542 (2006).
27. D. Lazinski, E. Grzadzielska, A. Das, *Cell*, **59**, 207 (1989).
28. S. Chattopadhyay, J. Garcia-Mena, J. DeVito, K. Wolska, A. Das, *Proc. Nat. Acad. Sci. USA* **92**, 4061 (1995).
29. P. Legault, J. Li, J. Mogridge, L. E. Kay, J. Greenblatt, *Cell* **93**, 289 (1998).
30. M. Scharpf *et al.*, *Eur. J. Biochem.* **267**, 2397 (2000).
31. S. J. Greive, A. F. Lins, P. H. von Hippel, *J. Biol. Chem.* **280**, 36397 (2005).
32. G. A. Belogurov, R. A. Mooney, V. Svetlov, R. Landick, I. Artsimovitch, *EMBO J.* **28**, 112 (2009).
33. G. A. Belogurov *et al.*, *Mol Cell* **26**, 117 (2007).
34. T. M. Arthur, L. C. Anthony, R. R. Burgess, *J. Biol. Chem.* **275**, 23113 (2000).
35. A. Sevostyanova, V. Svetlov, D. G. Vassilyev, I. Artsimovitch, *Proc. Nat. Acad. Sci. USA* **105**, 865 (2008).
36. R. A. Mooney, S. A. Darst, R. Landick, *Mol. Cell* **20**, 335 (2005).
37. B. T. Wimberly *et al.*, *Nature* **407**, 327 (2000).
38. N. C. Kyrpides, C. R. Woese, C. A. Ouzounis, *Trends Biochem. Sci.* **21**, 425 (1996).
39. T. Steiner, J. T. Kaiser, S. Marinkovic, R. Huber, M. C. Wahl, *EMBO J.* **21**, 4641 (2002).
40. B. S. Schuwirth *et al.*, *Science* **310**, 827 (2005).

41. T. M. Schmeing, V. Ramakrishnan, *Nature* **461**, 1234 (2009).

Danksagung

Die vorliegende Arbeit wurde von September 2006 bis Juli 2009 am Lehrstuhl Biopolymere und Forschungszentrum für Bio-Makromoleküle unter der Leitung von Prof. Dr. Paul Rösch erstellt. Ihm gilt mein besonderer Dank für die Überlassung eines interessanten und anspruchsvollen Themas, den ausgezeichneten Arbeitsmöglichkeiten und dem Interesse am Fortgang und Gelingen dieser Arbeit.

Dr. Kristian Schweimer möchte ich für die Unterstützung in allen Bereichen rund um die NMR-Spektroskopie danken. Bei Prof. Dr. Birgitta Wöhrl möchte ich mich für ihre Unterstützung und Beantwortung aller genetischen und das Labor betreffenden Fragen bedanken. Mein besonderer Dank gilt Ramona Heissmann für ihre erstklassige Unterstützung im Labor. Rainer Hofmann möchte ich für die Hilfe bei der Lösung von Computerproblemen danken. Gudrun Wagner und Violaine Zigan möchte ich für die Unterstützung in allen Verwaltungsangelegenheiten meinen Dank aussprechen.

Bedanken möchte ich mich außerdem bei Hanna Berkner, Andrea Hager, Maximilian Hartl, Nadine Herz, Claudia Knake, Berit Leo, Christian Mangels, Irena Matečko, Sophie Pleißner, Ulrike Persau, Sujatha Pagadala Santhanam, Stephan Schwarzinger, Philipp Weiglmeier, Sabine Wenzel und Britta Zimmermann für die ausgesprochen gute Laboratmosphäre und die stetige Hilfsbereitschaft.

Zusätzlich möchte ich mich auch noch bei Prof. Max E. Gottesman (Columbia University, USA) bei der Unterstützung bei diversen Manusripten und bei ihm und seiner Arbeitsgruppe für den abwechslungsreichen und interessanten Aufenthalt in seinem Labor bedanken.

Außerdem möchte ich mich auch bei Prof. Dr. Markus C. Wahl (Freie Universität Berlin) für die gute Zusammenarbeit und die Unterstützung bei diversen Manuskripten bedanken.

