Der Transkriptionskomplex und die Nus-Faktoren – Ein Netzwerk an Interaktionen

Dissertation

Zur Erlangung des Doktorgrades

der Fakultät für Biologie, Chemie und Geowissenschaften

an der Universität Bayreuth

Vorgelegt von

M. Sc. Biochemiker

Martin Strauß

aus Schweinfurt

Bayreuth, 2015

Die vorliegende Arbeit wurde von September 2011 bis April 2015 in Bayreuth am Lehrstuhl Biopolymere und Forschungszentrum für Bio-Makromoleküle unter der Leitung von Prof. Dr. Paul Rösch angefertigt.

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

Dissertation eingereicht am: 21.04.2015
Zulassung durch die Promotionskommission: 29.04.2015
Wissenschaftliches Kolloquium: 28.07.2015

Amtierender Dekan: Prof. Dr. Rhett Kempe

Prüfungsausschuss:
Prof. Dr. Paul Rösch (Erster Gutachter)
Prof. Dr. Clemens Steegborn (Zweiter Gutachter)
Prof. Dr. Dirk Schüler (Vorsitzender)
Prof. Dr. Thomas Scheibel

Inhaltsverzeichnis

Zusammenfassung

Die Synthese von Ribonukleinsäuren (RNAs) ist ein zentraler Prozess in allen Organismen, wobei die RNA-Polymerase (RNAP) hierbei das wichtigste Enzym ist. In Bakterien besteht diese aus den fünf Untereinheiten $\alpha_2\beta\beta'\omega$. Der Ablauf der Transkription kann in die Phasen Initiation, Elongation und Termination unterteilt werden, die alle stark reguliert sind und von zahlreichen Faktoren beeinflusst werden. Die *N utilization substances* (Nus) NusA, NusB, NusE und NusG spielen hierbei eine wichtige Rolle, indem sie beispielsweise die Elongationsrate der RNAP modulieren oder die RNAP in eine terminationsresistente Form überführen. Die molekulare Basis der Interaktion der Nus-Faktoren mit der RNAP sowie mögliche Wechselwirkungen der Nus-Faktoren untereinander sind zu großen Teilen unbekannt, jedoch essentiell für ein umfassendes Verständnis der Transkriptionsregulation.

NusG ist als einziger Transkriptionsfaktor in Bakterien, Archaeen und Eukaryoten konserviert und hat eine Vielzahl von Aufgaben bei der Genexpression. Er besteht aus einer N- und einer C-terminalen Domäne (NTD und CTD), die in vielen Organismen flexibel miteinander verknüpft sind. NusG aus dem Humanpathogen *Mycobacterium tuberculosis* (*Mt*NusG) zeigt einige regulatorische Unterschiede zu NusG aus *Escherichia coli* (*Ec*NusG), da *Mt*NusG beispielsweise die intrinsische Termination stimuliert, während *Ec*NusG keinen Einfluss darauf hat. In dieser Arbeit wurde die Struktur der *Mt*NusG-CTD in Lösung bestimmt und ein Modell der *Mt*NusG-NTD generiert. Insgesamt ähnelt *Mt*NusG strukturell stark *Ec*NusG und wie bei diesem interagieren NTD und CTD nicht miteinander, was mittels Kernspinresonanz (NMR)-Spektroskopie demonstriert wurde. Im Vergleich zu *Ec*NusG ist allerdings sowohl der Aminoterminus als auch der Linker zwischen den Domänen verlängert. Da diese Bereiche aber nicht für die funktionellen Unterschiede verantwortlich sind, ist vermutlich die Bindung von *Mt*NusG an die RNAP anders als bei *Ec*NusG.

NusA besteht aus sechs Domänen und beeinflusst insbesondere die Elongation und die Termination. In dieser Arbeit wurde erstmals eine spezifische, direkte Interaktion zwischen NusG-NTD und der C-terminalen Domäne von NusA, NusA-AR2, über NMR-Spektroskopie und *pulldown*-Assays nachgewiesen. *In vitro*-Transkriptionstests ergaben weiterhin, dass NusA und NusG gemeinsam eine Pause induzieren können, die ein Faktor alleine nicht hervorruft. Zusammen mit den Ergebnissen von NMR-Verdrängungsexperimenten deuten die Daten darauf hin, dass die Interaktion bei der Rekrutierung von NusG an die RNAP, bei der Synchronisation von Transkription und Translation oder bei der Regulation der Terminationseffizienz eine Rolle spielt.

NMR-Spektroskopie eignet sich insbesondere zur Untersuchung von Dynamik und schwachen Interaktionen. Diese Prozesse sind bei der RNAP-Regulation essentiell. Da die RNAP aber aufgrund ihrer großen Molekülmasse nicht mit konventionellen NMR-Experimenten analysiert werden kann, wurde im Rahmen dieser Arbeit zunächst ein effizientes Protokoll etabliert, um aktive RNAP aus individuell exprimierten Untereinheiten *in vitro* zu assemblieren und zu reinigen. Dies erlaubte dann

die selektive Markierung der Methylgruppen von Ile, Leu und Val-Resten einer bestimmten Untereinheit mit [^{1}H, ^{13}C], während die restliche Untereinheit und alle anderen Untereinheiten deuteriert vorlagen (Methylgruppenmarkierung). Somit konnten [^{1}H, ^{13}C]-Korrelationsspektren der β'-Untereinheit im Gesamtenzym aufgenommen werden. Außerdem wurden alle RNAP-Untereinheiten einzeln in löslicher, funktionaler Form gereinigt und NMR-Experimente etabliert, um die RNAP-Untereinheit zu identifizieren mit der ein bestimmter Transkriptionsfaktor interagiert. Hierdurch wurden die vorgeschlagenen Bindungsstellen von NusG und NusA bestätigt und gezeigt, dass NusE direkt an die β-Untereinheit bindet. Dieser Ansatz lässt sich generell auf kleine bis mittelgroße RNAP-bindende Proteine oder kleine organische Verbindungen, beispielsweise Antibiotika, anwenden.

Schließlich wurde eine NMR-spektroskopische Methode entwickelt, um die RNAP-Bindungsfläche der Nus-Faktoren zu bestimmen. Hierfür wurde der methylgruppenmarkierte Faktor mit protonierter RNAP titriert. Nachdem der Ansatz mit NusG-NTD validiert worden war, wurde die RNAP-Bindungsstelle von NusA-NTD identifiziert. Dies erlaubte die Erstellung eines detaillierten Modells, wie NusA-NTD an die RNAP bindet und sich die naszierende RNA um NusA wickelt. Weiterhin wurde gezeigt, dass die Region, mit der NusE an die RNAP bindet, mit derjenigen überlappt, die für die Interaktion von NusE mit NusG-CTD verantwortlich ist. NMR-Verdrängungsexperimente ergaben, dass die Affinitäten von RNAP und NusG an NusE ähnlich sind, was darauf hindeutet, dass die Bindung von NusE an die RNAP bei der Antitermination wichtig sein könnte. Der methodische Ansatz zur Bestimmung der RNAP-Bindungsfläche kann allgemein auf Systeme übertragen werden, bei denen ein supramolekularer Komplex an einen kleinen Partner bindet.

Summary

The synthesis of ribonucleic acids (RNAs) is a central process in all organisms, with RNA polymerase (RNAP) being the key enzyme. RNAP consists of the five subunits $\alpha_2\beta\beta'\omega$ in bacteria. Transcription can be divided into the phases initiation, elongation, and termination, which are all highly regulated. RNAP is controlled by multiple factors, for example the N utilization substances (Nus) NusA, NusB, NusE, and NusG, which play an important role in modulating the RNAP elongation rate or in converting the RNAP into a termination resistant form. The molecular basis for the interaction of Nus factors with RNAP as well as possible, mutual interactions between Nus factors are mostly unknown. Their knowledge, however, is essential for the complete understanding of transcription regulation.

NusG is the only transcription factor that is conserved in bacteria, archaea and eukaryotes, having a variety of functions in gene expression. It consists of an N- and a C-terminal domain (NTD and CTD), which are flexibly connected in most organisms. NusG from the human pathogen *Mycobacterium tuberculosis* (*Mt*NusG) shows some regulatory differences to NusG from *Escherichia coli* (*Ec*NusG), as for example, *Mt*NusG stimulates intrinsic termination, while *Ec*NusG has no influence on it. In this work the solution structure of *Mt*NusG-CTD was determined and a model of *Mt*NusG-NTD was generated. Altogether the structures of *Mt*NusG and *Ec*NusG are highly similar and just like in *Ec*NusG the NTD and CTD of *Mt*NusG do not interact as demonstrated by nuclear magnetic resonance (NMR) spectroscopy. The amino terminus and the linker between the two domains are, however, elongated in *Mt*NusG, but as these regions are not responsible for the functional differences, the interaction with RNAP might be different for *Mt*NusG and *Ec*NusG.

NusA consists of six domains and affects particularly elongation and termination. In this work a direct, specific interaction between NusG-NTD and the C-terminal NusA domain (NusA-AR2) has been demonstrated, using NMR spectroscopy and pull-down assays. Furthermore *in vitro* transcription assays showed, that NusA and NusG together are able to induce a novel pause, which is not evoked by one factor alone. Together with the results of NMR displacement experiments these data suggest that the interaction plays an important role in the recruitment of NusG to RNAP, the synchronization of transcription and translation or the regulation of the termination efficiency.

NMR spectroscopy is a technique that is especially convenient for the investigation of dynamics and weak interactions, processes essential for RNAP regulation. Due to its high molecular mass, however, RNAP cannot be studied using conventional NMR experiments. Thus, in this work, an efficient protocol was established for the *in vitro* assembly of active RNAP from its individually expressed subunits and its purification. This allowed the selective labeling of methyl groups of Ile, Leu and Val residues of a certain subunit with [^{1}H, ^{13}C], while the remaining subunit and all other subunits were deuterated (methyl group labeling). Using this approach [^{1}H, ^{13}C] correlation spectra of the β' subunit

within the complete RNAP were recorded. Furthermore, all RNAP subunits were purified individually in a soluble form and NMR experiments were established in order to identify the RNAP subunit a certain transcription factor interacts with. Hereby the proposed binding sites for NusG and NusA were confirmed and it was shown that NusE directly binds to the β subunit. The method is generally applicable to other small or medium sized RNAP binding proteins or small organic compounds, like antibiotics.

Finally, an NMR spectroscopic method was developed to determine the RNAP binding surfaces of Nus factors by titrating the methyl group-labeled factor with protonated RNAP. Having validated this approach with NusG-NTD, the RNAP binding surface of NusA-NTD was identified. This enabled the generation of a detailed model of how NusA-NTD binds to RNAP and how the nascent RNA is wrapped around NusA. I also found that the NusE region which binds to RNAP overlaps with the one involved into the interaction of NusE with NusG-CTD. NMR displacement measurements yielded similar affinities of RNAP and NusG for NusE, suggesting that the binding of NusE to RNAP might be important in antitermination. This approach to determine the RNAP binding surface can be generalized and is transferable to other systems, in which a supramolecular complex binds to a small partner.

1. Einleitung

1.1 Bakterielle Transkription

1.1.1 Die RNA-Polymerase

Die Transkription von Abschnitten der Desoxyribonukleinsäure (DNA) in Ribonukleinsäure (RNA) ist der erste Schritt zur Genexpression und daher für alle Organismen von zentraler Bedeutung. Das hierfür benötigte Enzym ist die RNA-Polymerase (RNAP). Die durch die RNAP erstellten RNAs können nicht nur als Vorlage zur Proteinbiosynthese verwendet werden (*messenger* RNA, mRNA), sondern können selbst das finale Genprodukt sein und bestimmte Aufgaben in der Zelle übernehmen. Transfer-RNAs (tRNAs) und ribosomale RNAs (rRNAs) ermöglichen beispielsweise den Transport von Aminosäuren zum Ribosom, bzw. sind am Aufbau des Ribosoms beteiligt. Des Weiteren können nichtkodierende RNA-Stränge die Transkription regulieren oder die mRNA-Stabilität und -Translation beeinflussen (zusammengefasst in Storz *et al.*, 2006).

Während die RNAPs in Mitochondrien, Chloroplasten und Bakteriophagen aus einer Untereinheit bestehen, verwenden alle Lebewesen für die zelluläre Transkription aus mehreren Untereinheiten aufgebaute RNAPs (Gaspari *et al.*, 2004; Werner und Grohmann, 2011). Alle diese RNAPs ähneln sich hinsichtlich der Interaktion mit Nukleinsäuren, der Struktur und dem Katalysemechanismus. Während archaeelle und eukaryotische RNAPs aus 12-17 Polypeptidketten bestehen, sind die bakteriellen RNAPs mit fünf Untereinheiten ($\alpha_2\beta\beta'\omega$, Kern-RNAP) einfacher aufgebaut (Abb. 1, Ebright, 2000).

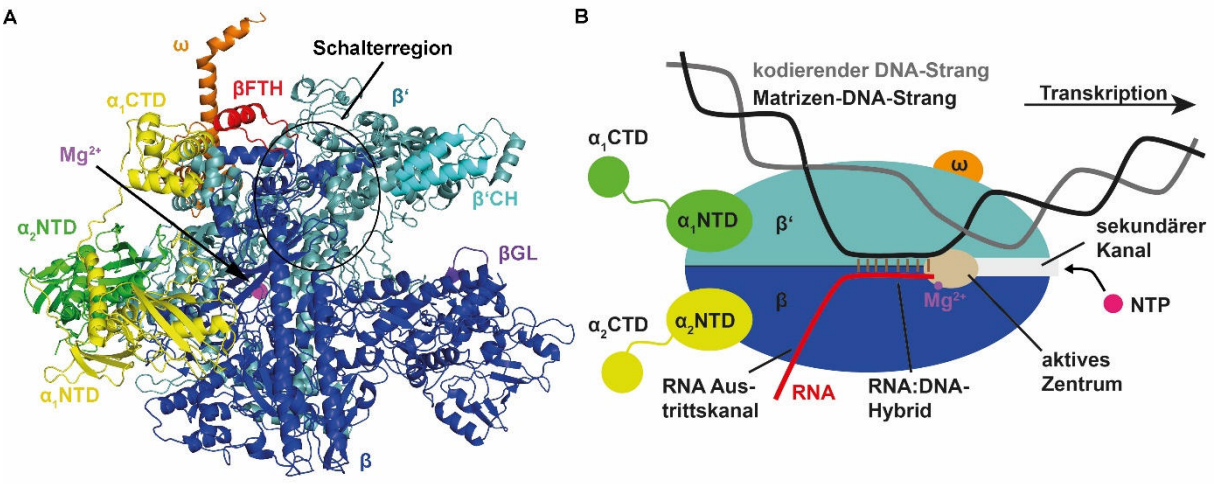


Abbildung 1: Struktur der bakteriellen RNAP. A) Die Proteinstruktur der *Escherichia coli* (*E. coli*) RNAP ist in Cartoondarstellung gezeigt (Proteindatenbank- (PDB-) Code: 4JKR). Die RNAP besteht aus zwei α-Untereinheiten (gelb und grün) und einer β- (dunkelblau), β‘-(dunkeltürkis) und ω-Untereinheit (orange). Zusätzlich ist das permanent gebundene Mg^{2+}-Ion eingezeichnet (lila Kugel). Die Schalterregion ist durch einen schwarzen Kreis gekennzeichnet und regulatorisch wichtige Elemente sind farblich hervorgehoben; β-Pfortenschleife (β *gate loop*, βGL, lila); β'-Klammerhelices (β‘ *clamp helices*, β'CH, hellblau) und β Klappenspitzehelix (β *flap tip helix*, βFTH, rot). **B)** Schematische Darstellung der elongierenden RNAP. Das Farbschema ist analog zu A) und wichtige Bereiche der RNAP sind beschriftet.

Die Struktur der RNAP wird häufig mit einer Krebsschere verglichen, in deren zentralen Spalte die DNA gebunden wird und die RNA-Synthese stattfindet (Zhang *et al.*, 1999). Die beiden Scheren bestehen aus der β- und der β'-Untereinheit, welche das aktive Zentrum ausbilden und zwei Mg^{2+}-Ionen koordinieren (Zhang *et al.*, 1999). Das erste Magnesiumion ist durch drei Aspartatreste fest am aktiven Zentrum positioniert, während das zweite zusammen mit dem entsprechenden Nukleotid zum aktiven Zentrum gebracht wird (Zaychikov *et al.*, 1996; Sosunov *et al.*, 2003). Die α-Untereinheiten sind nicht direkt Teil des aktiven Zentrums, sind jedoch in engem Kontakt mit der β- bzw. β'-Untereinheit. Die Dimerisierung der beiden α-Untereinheiten ist vermutlich der erste Schritt beim Zusammenbau der RNAP (Ishihama, 1981; Wang *et al.*, 1997). Die carboxyterminalen Domänen (αCTDs) sind flexibel mit dem Rest der RNAP verbunden und nehmen regulatorische Funktionen wahr (Ebright und Busby, 1995; Ishihama, 1992; Lee *et al.*, 2012;). Ihre Struktur konnte isoliert von der übrigen RNAP durch Kernspinresonanzspektroskopie (NMR-Spektroskopie) gelöst werden (Jeon *et al.*, 1995). Nur in einer 2013 veröffentlichten *Escherichia coli* (*E. coli*) RNAP-Kristallstruktur ist eine der beiden αCTDs aufgelöst (Murakami, 2013). Hier ist der Linker zwischen der αCTD und der αNTD vollkommen gestreckt und die αCTD ist nahe der ω-Untereinheit an die RNAP gebunden. Die ω-Untereinheit ist für die Funktion der RNAP nicht essentiell, sondern sie unterstützt vermutlich die korrekte Assemblierung der Kern-RNAP und die Bindung des σ-Faktors (Ghosh *et al.*, 2001; Mustaev *et al.*, 1997). Bei den RNAP-Strukturen von *Thermus aquaticus* und *Thermus thermophilus* (*T. thermophilus*) ist die ω-Untereinheit um den C-Terminus der β'-Untereinheit gewickelt, während diese Interaktion bei *E. coli* nicht vorhanden ist (Zhang *et al.*, 1999; Murakami, 2013; Murakami *et al.*, 2002).

Es konnte in Kristallstrukturen und in Lösung gezeigt werden, dass die krebsscherenartige Struktur der RNAP eine offene oder eine geschlossene Konformation einnehmen kann. Diese unterscheiden sich durch eine Drehbewegung in der Schalter-Region, nahe der Basis der beiden Scheren (Zhang *et al.*, 1999; Cramer *et al.*, 2001). Im offenen Zustand der RNAP kann die doppelsträngige DNA zum aktiven Zentrum gebracht und entwunden werden. Die geschlossene Form ist wichtig für eine starke DNA-Bindung und damit für die hohe Stabilität und Prozessivität während der Elongation.

Weitere wichtige strukturelle Merkmale der RNAP sind der RNA-Austrittskanal, durch welchen die naszierende RNA nach außen geführt wird und der sekundäre Kanal, durch welchen vermutlich die Ribonukleosidtriphosphate (NTPs) zum aktiven Zentrum gebracht werden (Zhang *et al.*, 1999; Cramer *et al.*, 2000; Korzheva *et al.*, 2000). Die β-Pfortenschleife (*β gate loop,* βGL) sorgt dafür, dass keine doppelsträngige DNA gebunden wird und spielt zusammen mit den β'-Klammerhelices (*β' clamp helices*, β'CH) eine wichtige Rolle bei der Regulation. Die β-Klappenspitzenhelix (β *flap tip helix*, βFTH) befindet sich an der Oberfläche der RNAP in der Nähe des RNA-Austrittskanals und ist eine Bindungsstelle für regulatorische Proteine (Vassylyev *et al.*, 2002).

1.1.2 Ablauf der bakteriellen Transkription

Analog zur Replikation und Translation kann die Transkription in die drei Abschnitte Initiation, Elongation und Termination eingeteilt werden. Eine schematische Darstellung des Transkriptionszyklus ist in Abbildung 2 zu sehen.

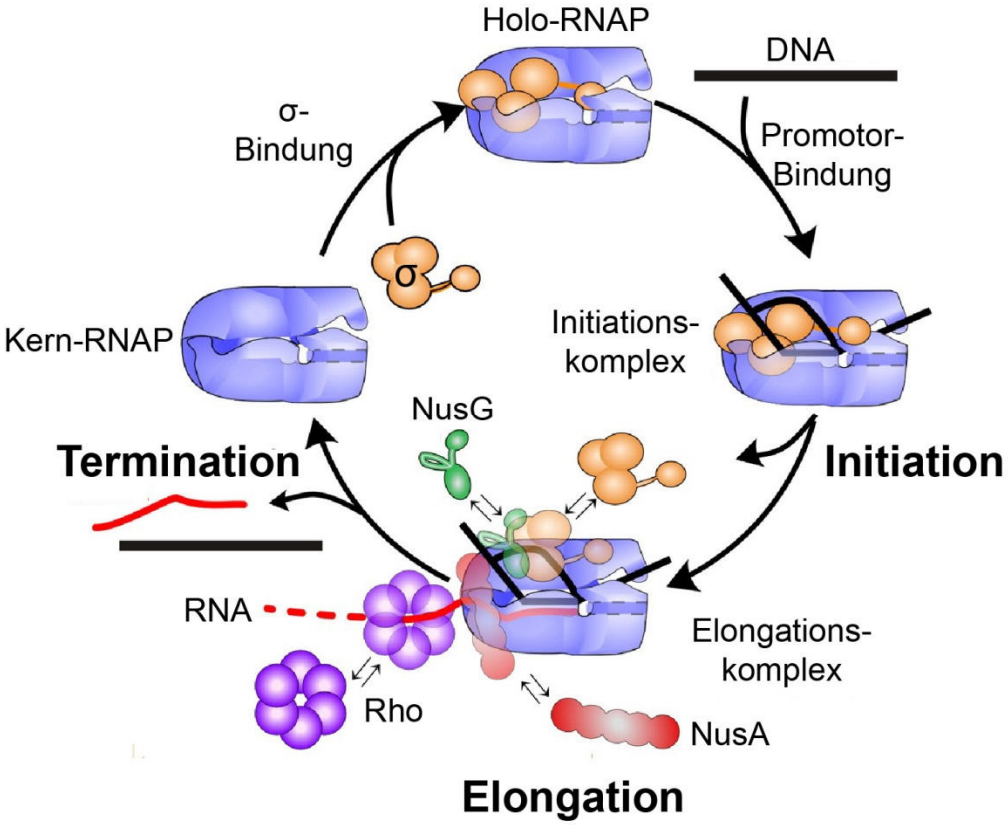


Abbildung 2: Schematische Darstellung des Transkriptionszyklus. Gezeigt sind die drei Phasen der Transkription: Initiation, Elongation und Termination sowie der Einfluss verschiedener Regulatoren. Zu Beginn bindet die Kern-RNAP den σ-Faktor, wodurch sich das Holoenzym bildet und Promotorregionen auf der DNA spezifisch erkannt werden können. Dies führt zur Ausbildung des Initiationskomplexes. Zu Beginn der Elongationsphase geht der starke Kontakt zum σ-Faktor verloren und die RNAP synthetisiert Nukleotide an die naszierende RNA. Dieser Vorgang ist nicht kontinuierlich sondern hängt von der jeweiligen Matrize sowie weiteren Faktoren, beispielsweise NusA (rot) und NusG (grün), ab. An bestimmten Stellen kommt es zur Termination, wobei die synthetisierte RNA aus der Polymerase entlassen wird und die RNAP wieder als Kernenzym vorliegt. Weitere Details sind dem Text zu entnehmen. (Abbildung verändert nach Mooney *et al.*, 2009a.)

1.1.2.1 Initiation

Zu Beginn der Transkription liegt die RNAP als Kernenzym ($\alpha_2\beta\beta'\omega$) vor. In dieser Konformation bindet sie DNA unspezifisch und kann an dieser entlanggleiten (Sakata-Sogawa und Shimamoto, 2004). Für die Initiation wird ein weiteres Protein, der bakterielle Transkriptionsinitiations- oder σ-Faktor, gebunden (zusammengefasst in deHaseth und Helmann, 1995; deHaseth *et al.*, 1998). Der hieraus entstandene Komplex wird als Holoenzym bezeichnet und ist in der Lage Promotorregionen auf der DNA zu erkennen. Hierdurch bildet sich der geschlossene Promotorkomplex, in welchem die DNA noch doppelsträngig vorliegt (Burgess und Anthony, 2001; Borukhov und Severinov, 2002). In *E. coli* sind mehrere unterschiedliche σ-Faktoren bekannt, welche verschiedene Promotorregionen erkennen können (zusammengefasst in Paget und Helmann, 2003). In logarithmisch wachsenden

Zellen ist der Haushalts-σ-Faktor, σ^{70}, am häufigsten zu finden (deHaseth und Helmann, 1995). Dieser erkennt vor allem DNA-Konsensussequenzen in der -10- und der -35-Region. Die Zahl bezieht sich hierbei auf die Position relativ zum Transkriptionsinitiationspunkt. Um die Transkription zu starten interagiert der σ-Faktor mit der Promotorregion und induziert eine konformationelle Änderung der RNAP, welche zur starken Bindung der DNA und zum Aufschmelzen der doppelsträngigen DNA im Bereich -11 bis +4 führt (zusammengefasst in Haugen *et al.*, 2008). Der Mechanismus für das Auftrennen des Doppelstranges ist im Detail nicht bekannt, es erfolgt jedoch in Abhängigkeit von Mg^{2+} (Suh *et al.*, 1992; Zaychikov *et al.*, 1997). Da die DNA nun abschnittsweise einzelsträngig vorliegt wird der Komplex als offener Promotorkomplex bezeichnet. Der σ-Faktor bindet anschließend an die -10 und an benachbarte Regionen (erweiterte -10-Region), wodurch der Matrizenstrang im aktiven Zentrum der RNAP positioniert wird (Guo *et al.*, 2000; Barne *et al.*, 1997; Bown *et al.*, 1999).
Im Gegensatz zur DNA-Polymerase benötigt die RNAP keinen Primer zum Start der Polymerisation. Stattdessen werden die ersten beiden NTPs zeitgleich an die Matrizen-DNA gebracht und die erste Phosphodiesterbindung synthetisiert (Basu *et al.*, 2014). Die RNAP beginnt nun damit, immer wieder kurze RNA-Stücke zu synthetisieren und freizusetzen. Dieser Prozess wird als „abortive Initiation“ bezeichnet (Carpousis und Gralla, 1980; Vo *et al.*, 2003). Die Funktion der abortiven Initiation ist nicht eindeutig geklärt. Vermutlich findet hierdurch ein Korrekturlesen des Promotors statt, wodurch die Genexpression reguliert wird (Liu *et al.*, 2011). Hierbei wird davon ausgegangen, dass die RNAP den Promotor nur verlassen kann, falls generelle und promotorspezifische Interaktionspartner vorhanden sind. Grundlage für diese Annahme war die Beobachtung, dass kurze RNA-Stränge mit einer Länge von weniger als fünf Nukleotiden in Kristallstrukturen stark verformt vorliegen und das DNA:RNA-Hybrid energetisch ungünstige Konformationen einnimmt. Sobald alle Transkriptionsfaktoren vorhanden sind, kann die RNAP schnell längere RNA-Stücke synthetisieren und der Komplex wird stabilisiert (Liu *et al.*, 2011).
Während der Initiation können die RNAP αCTDs auf zwei Arten eine regulatorische Funktion ausüben. Zum einen kann eine stromaufwärts des Promotors gelegene *cis*-aktive DNA-Sequenz (*upstream promotor element*) gebunden werden, was zu einer Erhöhung der Promotoraktivität führt (Ross *et al.*, 1993). Zum anderen konnte gezeigt werden, dass Transaktivatoren, wie das Katabolitgen-Aktivatorprotein (CAP) und das oxidative Stressregulatorprotein (OxyR) gebunden werden (Tao *et al.*, 1993; Benoff *et al.*, 2002).

1.1.2.2 Elongation

Beim Übergang vom Initiations- zum Elongationskomplex müssen die spezifischen Interaktionen mit dem Promotor aufgelöst werden. Die RNAP zieht hierzu stromabwärts liegende DNA in sich hinein und diese zusammengedrückte DNA-Struktur stellt genügend Energie zur Verfügung, um die Kontakte zwischen dem σ-Faktor und dem Promotor aufzulösen (Kapanidis *et al.*, 2006; Revyakin *et al.*, 2006). Der σ-Faktor dissoziiert zu Beginn der Elongationsphase häufig von der RNAP ab, kann jedoch auch an der RNAP gebunden bleiben (Mooney *et al.*, 2009a; Shimamoto *et al.*, 1986;

Mukhopadhyay *et al.*, 2001; Bar-Nahum und Nudler, 2001). Während der Elongationsphase transloziert die RNAP entlang dem DNA-Matrizenstrang und synthetisiert komplementär dazu basenweise einen neuen RNA-Strang (Abbondanzieri *et al.*, 2005). Die Synthesegeschwindigkeit variiert stark zwischen 10-135 Nukleotiden pro Sekunde, da sie abhängig von den Wachstumsbedingungen und dem jeweiligen Transkript ist (Gotta *et al.*, 1991; Vogel und Jensen, 1994; Condon *et al.*, 1993).

Während der Elongation liegt der Matrizenstrang in einer um ca. 90° geknickten Konformation vor. Der Knick befindet sich hierbei im aktiven Zentrum (Wang *et al.*, 2006; Vassylyev *et al.*, 2007a). Die naszierende RNA wird durch den RNA-Austrittskanal an die Außenseite der RNAP gebracht, wobei kaum eine Stabilisierung durch die RNAP stattfindet (Vassylyev *et al.*, 2007a). Das RNA:DNA-Hybrid ist 8-9 Basenpaare (bp) lang und befindet sich im aktiven Zentrum der RNAP. Zwischen dem Phosphatrückgrat des Hybrids und der RNAP werden hauptsächlich polare und van-der-Waals-Kontakte ausbildet, wodurch die Transkription nicht durch zu starke spezifische Wechselwirkungen verlangsamt wird (Vassylyev *et al.*, 2007a).

Der Einbau eines Nukleotids besteht aus den Schritten NTP-Bindung, Katalyse der RNA-Elongation, Entfernung des Pyrophosphats (PP_i) und Translokation (Abb. 3, Zhang und Landick, 2009). Zu Beginn des Zyklus liegt die RNAP im posttranslozierten Zustand vor. Das 3'-Ende der wachsenden RNA-Kette ist hierbei an der Produktstelle positioniert (i-1), während das einzubauende NTP an der Insertionsstelle (i+1) eingebaut wird. Die DNA liegt lediglich am aktiven Zentrum (i+1) einzelsträngig vor, sodass immer nur ein neues NTP eingebaut werden kann (Vassylyev *et al.*, 2007a; Cramer, 2007).

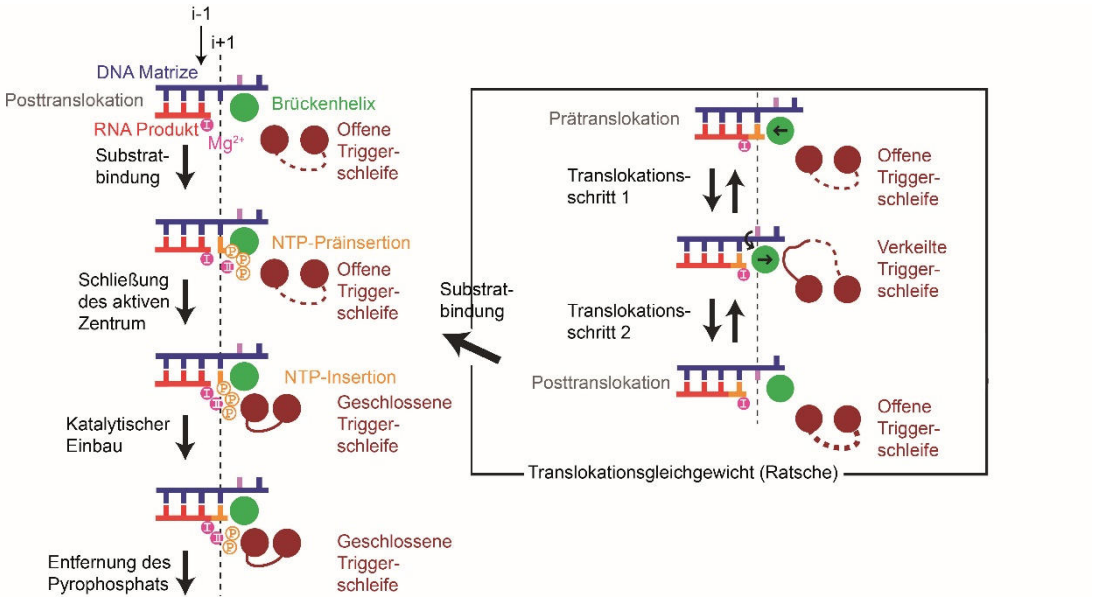


Abbildung 3: Schematische Darstellung der Transkriptionselongation. Gezeigt ist das DNA:RNA-Hybrid nach der Translokation mit einer freien Bindungsstelle an der i+1 Position (gestrichelte Linie). Der Einbau des nächsten Nukleotids erfolgt durch magnesiumabhängige Katalyse, wobei die Triggerschleife von einer offenen, unstrukturierten Form in eine geschlossene α-helikale Form übergeht um das einzubauende Nukleotid richtig zu positionieren. Nach Entfernung des Pyrophosphats erfolgt die Translokation, sodass erneut eine Base des DNA-Matrizenstrangs frei vorliegt. Weitere Details sind dem Text zu entnehmen. (Abbildung verändert nach Brückner und Cramer, 2008.)

Die Triggerschleife spielt bei diesem Prozess eine zentrale Rolle (Abb. 4A). Sie liegt ungefaltet vor und hilft, neben einigen Resten des sekundären Kanals, bei der Erkennung des richtigen NTPs und verhindert den Einbau von Desoxyribonukleosidtriphosphaten (dNTPs, Wang *et al.*, 2006; Holmes *et al.*, 2006; Kaplan *et al.*, 2008; Kireeva *et al.*, 2008). Das neue NTP wird zunächst in der Nähe des aktiven Zentrums gebunden, wobei der sekundäre Kanal in einer offenen Form vorliegt und die Triggerschleife immer noch unstrukturiert ist (Vassylyev *et al.*, 2007b). Dies wird als Präinsertionszustand bezeichnet. Im nächsten Reaktionsschritt bildet die Triggerschleife eine Helix (Triggerhelix) aus. Hierdurch wird die Größe des sekundären Kanals reduziert und das NTP optimal positioniert, sodass es nicht dissoziieren kann (Vassylyev *et al.*, 2007b). Die Elongation erfolgt dann durch einen Mg^{2+}-abhängigen, S_N2 nukleophilen Angriff der RNA-3'-Hydroxylgruppe auf das α-Phosphoratom des NTP, wobei PP_i abgespalten wird (Yee *et al.*, 1979; Tomar und Artsimovitch, 2013). Vermutlich führt das Entfernen des PP_i zu einer keilförmigen Konformation der Triggerhelix, wodurch die Brückenhelix ebenfalls ihre Konformation ändert und die Translokation stattfindet (Abb. 4B, Fouqueau *et al.*, 2013; Feig und Burton, 2010). Anschließend liegt wieder eine ungepaarte Base der DNA im aktiven Zentrum vor und das nächste NTP kann eingebaut werden. Dies geschieht bis zur Transkriptionstermination. Da die Vorwärtsbewegung ohne zusätzliche Energie abläuft und der prä- und posttranslozierte Zustand im Gleichgewicht vorliegen, spricht man bei dem Mechanismus von einer „Brownschen Ratsche". Aufgrund der strukturellen Änderungen der Triggerschleife von unstrukturiert zu helikal bzw. keilartig und dem Einbau des nächsten NTPs ist die Translokation begünstigt, wodurch ein Zurückrutschen der RNAP verhindert wird (Brückner und Cramer, 2008; Tagami *et al.*, 2010; Bar-Nahum *et al.*, 2005).

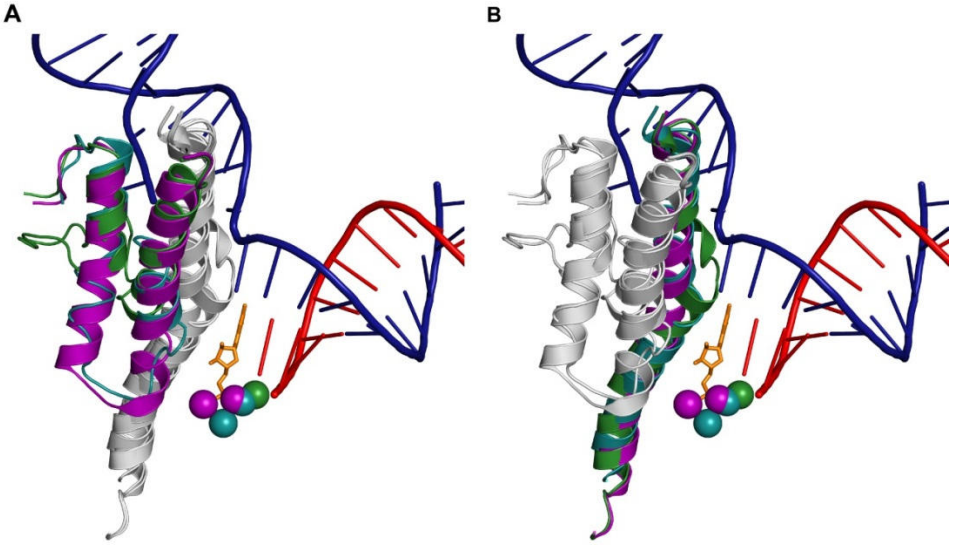


Abbildung 4: Strukturelle Änderungen der Triggerschleife (A) und der Brückenhelix (B) im Zentrum der RNAP während der Transkriptionselongation. Gezeigt ist die DNA in blau, der neu synthetisierte RNA-Strang in rot (aus PDB-Code: 1IW7, Cartoondarstellung). **A)** Struktur der Triggerschleife im offenen (türkis, PDB-Code: 2E2H, Wang *et al.*, 2006), geschlossenen (lila, PDB-Code: 2O5J, Vassylyev *et al.*, 2007b) und verkeilten (grün, PDB-Code: 1IW7, Vassylyev *et al.*, 2002) Zustand. Die katalytischen Magnesiumionen sind analog eingefärbt. Die Brückenhelix aller Zustände ist in grau, das NTP im aktiven Zentrum ist in orange gezeigt. **B)** Die strukturellen Zustände der Brückenhelix sind farblich analog zu A) eingezeichnet. Die Postitionen der Triggerschleife sind in grau dargestellt.

1.1.2.3 Termination

Die Transkriptionstermination hat in der Zelle zwei wichtige Funktionen. Zum einen werden benachbarte Transkriptionseinheiten reguliert, zum anderen kann sie als regulatorisches Element innerhalb eines Gens eingesetzt werden. In der Terminationsphase wird die RNA-Synthese beendet, das DNA:RNA-Hybrid dissoziiert, die aufgeschmolzene Region der DNA wird wieder gewunden und die RNAP verlässt die Matrize. Obwohl der Transkriptionskomplex extrem stabil ist und *in vitro* eine Halbwertszeit von mehreren Tagen hat, kann die Termination an definierten Stellen abrupt stattfinden (Arndt und Chamberlin, 1990). Für die Transkriptionstermination sind in Bakterien zwei Mechanismen vorhanden, die intrinsische und die Rho-abhängige Termination.

Die intrinsische Termination

Der Hauptteil der Terminatoren in *E. coli* sind mit ca. 80 % intrinsische Terminatoren (Peters *et al.*, 2009). Für die intrinsische Termination sind keine weiteren Faktoren notwendig. Charakteristisch für intrinsische Terminationsstellen ist, dass das naszierende RNA-Transkript eine palindromische GC-reiche Sequenz enthält, welcher eine uridinreiche Region folgt (Platt, 1986; Brendel *et al.*, 1986). Durch die vielen schwachen A:U Basenpaare im Inneren der RNAP ist das RNA:DNA-Hybrid stark destabilisiert, wodurch die RNAP pausiert und die palindromische GC-reiche Region genügend Zeit hat eine Haarnadelschleife auszubilden (Gusarov und Nudler, 1999). Für den Mechanismus wie die Ausbildung der Haarnadelschleife letztlich zur Termination führt, sind drei Modelle vorhanden, wobei in allen die Destabilisierung des RNA:DNA-Hybrids eine Rolle spielt. Im allosterischen Modell interagiert die Terminationshaarnadelschleife mit strukturellen Elementen des RNA Austrittskanals der RNAP (Toulokhonov und Landick, 2003). Dies führt zu konformationellen Änderungen in der RNAP, der Destabilisierung des RNA:DNA-Hybrids und letztendlich zum Zusammenbruch der Transkriptionsblase (Toulokhonov *et al.*, 2001; Epshtein *et al.*, 2007). Beim Vorwärtstranslokations-modell wird durch die Entstehung der Haarnadelschleife die RNAP entlang der DNA gedrückt, wobei keine RNA-Elongation stattfindet. Die RNAP ist dann in einem hypertranslozierten Zustand mit einem verkleinerten, destabilisierten RNA:DNA-Hybrid, wodurch es zur Termination kommt (Santangelo und Roberts, 2004; Yarnell und Roberts, 1999). Im Scherenmodell wird durch die Bildung der Haarnadelschleife die RNA aus dem RNA:DNA-Hybriden herausgezogen, ohne dass eine Translokation der RNAP stattfindet (Toulokhonov und Landick, 2003; Macdonald *et al.*, 1993). Die drei Mechanismen treten abhängig vom spezifischen Terminator auf und es können auch mehrere Effekte gleichzeitig eine Rolle spielen (Larson *et al.*, 2008).

Die Rho-abhängige Termination

Im Gegensatz zur intrinsischen Termination wird bei der Rho-abhängigen Termination ein zusätzliches Protein, der Rho-Faktor (Rho), benötigt. Rho-abhängige Terminatoren können innerhalb eines Gens vorliegen (intragen) oder zwischen zwei Genen (intergen, Peters *et al.*, 2009). Sie spielen eine Rolle bei Genpolarität, wodurch stromaufwärts gelegene Gene häufiger transkribiert werden als

stromabwärts gelegene (Richardson *et al.*, 1975). Hierdurch wird auch die Transkription horizontal erlangter DNA, beispielsweise aus Prophagen, blockiert (Cardinale *et al.*, 2008). Außerdem ist Rho fähig schädliche Hybride aus doppelsträngiger DNA und naszierender RNA außerhalb der RNAP (R-loops) aufzulösen (Leela *et al.*, 2013).

Rho ist eine ATP-abhängige Translokase, welche an unstrukturierte und ribosomenfreie RNA bindet, sich dann in Richtung der RNAP bewegt und dort zur Termination führt (zusammengefasst in Ciampi, 2006). In *E. coli* ist Rho ein Homohexamer, wobei die sechs Untereinheiten eine ringähnliche Struktur einnehmen, mit einem zentralen Kanal. Jedes Protomer besteht aus einer NTD mit einer primären RNA-Bindungsstelle und einer CTD. Die sechs CTDs bilden die sekundäre RNA-Bindungsstelle im zentralen Kanal des Hexamers (Miwa *et al.*, 1995; Thomsen und Berger, 2009). In der Grenzfläche zwischen den benachbarten Untereinheiten findet die ATP-Hydrolyse statt, was zur Translokation der RNA führt (Thomsen und Berger, 2009). Für die Bindung von RNA ist keine Konsensussequenz bekannt, jedoch werden cytidinreiche Regionen mit einer Länge von 70-80 Nukleotiden bevorzugt (Morgan *et al.*, 1985; Zhu und von Hippel, 1998). Diese werden als Rho-Anwendungsstellen (*Rho utilization sites, rut*) bezeichnet. Um die RNA an der sekundären Bindungsstelle zu binden liegt Rho zunächst in einem offenen Zustand, ähnlich einem Federring, vor und nimmt nach der RNA-Bindung an der sekundären Bindungsstelle die ringförmige Konformation ein (Abb. 5A, Skordalakes und Berger, 2006; Canals *et al.*, 2010; Kim und Patel, 2001).

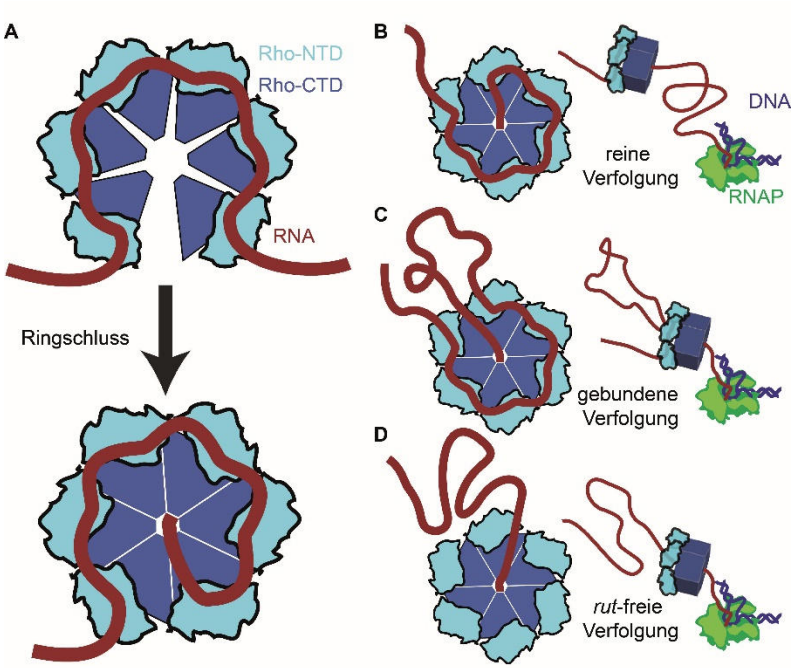


Abbildung 5: RNA-Bindung und Translokation durch Rho. A) Für die RNA-Bindung werden zunächst Kontakte mit der primären RNA-Bindungsstelle der Rho-NTDs ausgebildet. Anschließend wird die RNA an der sekundären RNA-Bindungsstelle gebunden und Rho bildet eine geschlossene ringartige Struktur aus. **B)** Beim reinen Verfolgungsmodell beruht die Translokation auf einem ATP/ADP-abhängigen Wechsel der RNA-Bindungsaffinität zwischen hoch- und niedrigaffin. **C)** Im gebundenen Verfolgungsmodell ist Rho durchgängig mittels der NTDs an der *rut* gebunden und fädelt die RNA durch die sekundäre Bindungsstelle. Dadurch kommt es zur Ausbildung einer RNA-Schlaufe. **D)** Beim *rut*-freien Verfolgungsmodell bindet Rho zunächst an die *rut* Bindestelle. Diese Interaktionen werden bei der Translokation allerdings aufgehoben. Weitere Details sind dem Text zu entnehmen. Die Einzelabbildungen stammen aus Koslover *et al.*, 2012 und wurden leicht bearbeitet.

Nach der RNA-Bindung transloziert Rho in Richtung der RNAP (5'→3'). Zunächst wurde angenommen, dass Rho die RNAP einholen kann, wenn diese verlangsamt oder gestoppt wird, und anschließend die Termination stattfindet (kinetische Kopplung, Jin *et al.*, 1992). Die Zusammenhänge zwischen Rho-abhängiger Termination und dem Pausieren der RNAP sind allerdings komplexer und können nicht nur hierauf zurückgeführt werden. Für die Translokation werden derzeit drei Modelle diskutiert (Abb. 5B-D). Im reinen Verfolgungsmodell wird angenommen, dass die NTDs abhängig vom Phosphorilierungszustand des gebundenen Adenosins (ATP oder ADP) immer wieder zwischen einem hoch- und niedrigaffinen Zustand umschalten. Durch das periodische Trennen und Binden an das Transkript findet demnach eine Translokation der RNA statt (Geiselmann *et al.*, 1993; Walstrom *et al.*, 1997). Dieses Modell ist unwahrscheinlich, da die ATP-Hydrolyse ausschließlich eng mit konformationellen Änderung in zwei Loopregionen der sekundären RNA Bindungsstelle zusammenhängt und für die Translokation die RNA an der sekundären Bindungsstelle gebunden sein muss (Miwa *et al.*, 1995; Thomsen und Berger, 2009; Wei und Richardson, 2001). Im gebundenen Verfolgungsmodell ist Rho durchgängig mittels der NTDs an der *rut* gebunden und fädelt die RNA durch die sekundäre Bindungsstelle. Dadurch kommt es zur Ausbildung einer RNA-Schlaufe (Steinmetz und Platt, 1994). Dieses Modell wurde durch Einzelmolekülkraftspektroskopie bestätigt (Koslover *et al.*, 2012). Beim *rut*-freien Verfolgungsmodell bindet Rho zunächst an die *rut* Bindestelle, transloziert dann aber entlang der RNA durch die Kontakte an der sekundären RNA-Bindestelle, wobei die anfänglichen Kontakte gebrochen werden.

Die Translokation in Richtung der RNAP spielt eine wichtige Rolle für die Rho-abhängige Termination. Die Dissoziation des Transkriptionskomplexes kann dabei bereits zehn Nukleotide stromabwärts der *rut* erfolgen oder 80-100 Nukleotide davon entfernt (Richardson und Richardson, 1996). Für die Termination werden ebenfalls drei Modelle diskutiert. Im ersten Modell transloziert Rho die RNA, wobei keine neuen Nukleotide an den RNA-Strang geknüpft werden. Das führt zu einer Destabilisierung des DNA:RNA-Hybrids in der RNAP und dadurch zur Termination (Richardson, 2002). Die Destabilisierung des DNA:RNA-Hybrids kann auch dadurch erzeugt werden, dass Rho die RNAP nach vorne drückt, wobei der RNA-Strang nicht verlängert wird (Park und Roberts, 2006). Beim dritten Mechanismus wird angenommen, dass Rho während der Transkription durchgehend an die RNAP und die naszierende RNA gebunden ist, wodurch sich eine Schlaufenstruktur ausbildet (Epshtein *et al.*, 2010). Durch die Translokation von Rho entlang der RNA wird die RNA-Schlaufe verkleinert und die ausgebildete Struktur ähnelt der einer Haarnadelschleife. Die Destabiliserung des DNA:RNA-Hybrids erfolgt dann ähnlich zu der intrinsischen Termination (Epshtein *et al.*, 2010). Dieses Modell wird kontrovers diskutiert. Durch Chromatin-Immunopreziptiations- (ChIP-chip-) Experimente ist bekannt, dass Rho während der Transkription immer in einem ähnlichen Verhältnis zur RNAP vorliegt (Mooney *et al.*, 2009a). Eine direkte oder starke Interaktion konnte jedoch weder mittels Einzelmolekülkraftspektroskopie noch in *Pulldowns* bestätigt werden (Koslover *et al.*, 2012; Kalyani *et al.*, 2011).

1.1.2.4 Antitermination

Terminatoren verhindern, dass benachbarte Gene abgelesen werden und können auch innerhalb eines Gens eine regulatorische Funktion einnehmen (Tomar und Artsimovitch, 2013; Santangelo und Artsimovitch, 2011). Um die stromabwärts gelegenen Gene transkribieren zu können, müssen Terminationssignale überlesen werden. Dieser Vorgang wird als Antitermination bezeichnet und kann durch Proteine, kleine Moleküle, RNAs oder auch die Temperatur geschehen. Eine Zusammenfassung einer Reihe an Antiterminatoren und ihrer Funktionsweise ist in Santangelo und Artsimovitch, 2011 zu finden. Der erste entdeckte Antiterminator war das N Protein des Bakteriophagen λ (λN). Der Bakteriophage λ kann sich im lysogenen Zustand in das *E. coli*-Genom integrieren oder im lytischen Zustand seine eigene DNA transkribieren und Phagenproteine herstellen, wodurch es zur Produktion neuer Phagen und zum Platzen des Wirts kommt (Oppenheim *et al.*, 2005). λN spielt im Lebenzyklus des Phagen eine wichtige Rolle, da es das Überlesen mehrerer Terminatoren ermöglicht und dadurch Proteine für den lytischen oder lysogenen Zyklus hergestellt werden können (Cheng *et al.*, 1995; Costantino *et al.*, 1990). Hierzu bindet λN es an die *nut*-RNA-Haarnadelschleife und rekrutiert die *N utilization substances* (Nus) NusA, NusB, NusE und NusG, um einen stabilen Antiterminationskomplex zu bilden (Mogridge *et al.*, 1995; Rees *et al.*, 1997; Das, 1993). Die Nus-Faktoren sind Proteine aus dem Wirt, spielen bei der Regulation der sieben ribosomalen RNA operons (*rrn*) eine Rolle und sind teilweise allgemeine Transkriptionsfaktoren.

1.2 Die Rolle der Nus-Faktoren in der Transkription

1.2.1 NusA

NusA aus *E. coli* besteht aus sechs Domänen (Abb. 6), hat eine Molekularmasse von 56 kDa und besitzt mehrere zelluläre Funktionen. Es verlangsamt die Transkriptionsgeschwindigkeit der RNAP durch die Verlängerung von Transkriptionspausen und führt neue Pausierungsstellen ein (Farnham *et al.*, 1982; Lau *et al.*, 1983; Kassavetis und Chamberlin, 1981). Diese Effekte sollten die intrinsische Termination begünstigen, da die Haarnadelstruktur mehr Zeit hat um sich auszubilden. NusA kann tatsächlich, beispielsweise bei dem Terminator λtR2, die Terminationseffizienz erhöhen (Schmidt und Chamberlin, 1987; Gusarov und Nudler, 2001). Bei anderen Terminatoren wie t*rpoB*a kann die Termination durch NusA aber auch abgeschwächt werden, indem NusA die RNA-Haarnadelschleife destabilisiert (Linn und Greenblatt, 1992; Ralling und Linn, 1987; Beuth *et al.*, 2005). Bei Rho-abhängigen Terminatoren ist die Rolle von NusA ebenfalls unklar und die Termination kann entweder inhibiert oder stimuliert werden (zusammengefasst in Borukhov *et al.*, 2005; Roberts *et al.*, 2008). NusA ist außerdem Teil des Antiterminationssystems (Mason und Greenblatt, 1991; Mason *et al.*, 1992b).

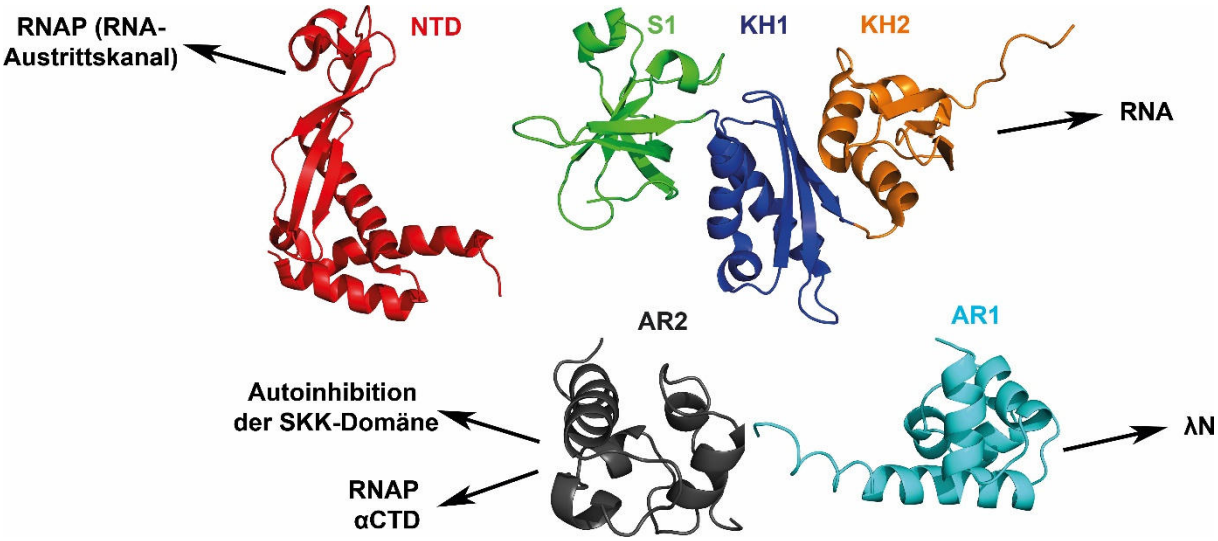


Abbildung 6: NusA-Struktur. Die Domänenstrukturen sind in Cartoondarstellung abgebildet. Die Pfeile zeigen an mit welchem Partner bzw. mit welchen Partnern die jeweilige Domäne interagieren kann. In *E. coli* besteht NusA aus einer NTD (rot, PDB-Code: 2KWP), drei RNA bindenden Domänen S1 (grün), KH1 (dunkelblau) und KH2 (orange), welche die SKK-Domäne ausbilden (PDB-Code: 1HH2, Worbs *et al.*, 2001), und zwei stark negativ geladenen C-terminalen Domänen AR1 und AR2 (hellblau und schwarz, PDB-Codes: 1WCL, 1WCN, Eisenmann *et al.*, 2004).

Der konservierte Teil von NusA besteht aus der NTD, welche durch eine Helix mit der RNA bindenden SKK-Domäne (S1, KH1 (K Homologie) und KH2) verbunden ist (Worbs *et al.*, 2001). Die NTD bindet nahe des RNA-Austrittskanals an die RNAP und obwohl bisher keine RNA-Bindung der NusA-NTD nachgewiesen wurde, wird angenommen, dass sie mit der naszierenden RNA interagiert und dadurch Transkriptionspausen verlängert (Yang *et al.*, 2009; Ha *et al.*, 2010; Mishra *et al.*, 2013). In *α*-, *β*- und *γ*-*Proteobacteria*, z. B. *E. coli* sowie in *Chlamydia* und *Treponema* sind zwei zusätzliche, stark negativ geladene carboxyterminale Domänen vorhanden: *acidic repeat* 1 (AR1) und AR2 (Mah *et al.*, 2000). Obwohl diese Domänen sich strukturell sehr ähnlich sind, führen sie unterschiedliche Funktionen aus. Die AR2-Domäne kann mit der SKK-Domäne interagieren und hat einen autoinhibitorischen Einfluss auf die RNA-Bindung (Mah *et al.*, 1999). Dieser wird aufgehoben, wenn AR2 an die αCTD der RNAP bindet (Mah *et al.*, 2000; Schweimer *et al.*, 2011). Die AR1-Domäne bindet den Antiterminator λN und unterstützt die λ-Antitermination (Bonin *et al.*, 2004; Eisenmann *et al.*, 2005; Prasch *et al.*, 2006). Es konnte allerdings gezeigt werden, dass diese Interaktion für die λN-abhängige Termination nicht essentiell ist (Mishra *et al.*, 2013).

1.2.2 NusE und NusB

NusE ist identisch mit dem ribosomalen Protein S10 und spielt sowohl bei der Transkription als auch bei der Translation eine wichtige Rolle (Friedman *et al.*, 1981). Bei der Translation übernimmt es, als Teil der 30S-Untereinheit des Ribosoms, eine strukturelle Funktion (Wimberly *et al.*, 2000). Während der Transkription ist NusE an die RNAP gebunden und wird vor allem bei der Antitermination benötigt (Mason und Greenblatt, 1991). Zusammen mit NusB kann NusE ein Heterodimer (NusB:NusE) ausbilden, welches an die RNA-Sequenzen *BoxA* und *BoxB* bindet (Abb. 7A, Stagno *et*

al., 2011; Mason *et al.*, 1992a; Luttgen *et al.*, 2002). *BoxA* und *BoxB* sind RNA-Kontrollsequenzen und sind sowohl in *nut*- als auch in *rrn*-Antiterminationssequenzen vorhanden (zusammengefasst in Morgan, 1986). Die Bindung von NusB:NusE an die *BoxA* ist der erste Schritt bei der Assemblierung des Antiterminationskomplexes (Stagno *et al.*, 2011; Greive *et al.*, 2005). Obwohl isoliertes NusB bereits an *BoxA* bindet, wird die Bindungsaffinität im Komplex aus NusB und NusE durch eine Vergrößerung der Bindungsfläche um das zehnfache gesteigert (Luttgen *et al.*, 2002; Greive *et al.*, 2005). Durch Überexpression von *nusE* in einer *nusB* Deletionsmutante konnte gezeigt werden, dass NusE die aktive Komponente im Dimer ist und NusB lediglich NusE zur *BoxA* rekrutiert (Luo *et al.*, 2008; Weisberg, 2008). Da die Bindungsfläche von NusE an NusB und an das Ribosom überlappen, kann NusE nicht Teil des Ribosoms sein während es an NusB gebunden ist (Luo *et al.*, 2008). NusE bindet außerdem als Teil des Ribosoms über NusG indirekt an die RNAP, wodurch es ein wichtiges Protein für die Kopplung der Transkription mit der Translation ist (Burmann *et al.*, 2010).

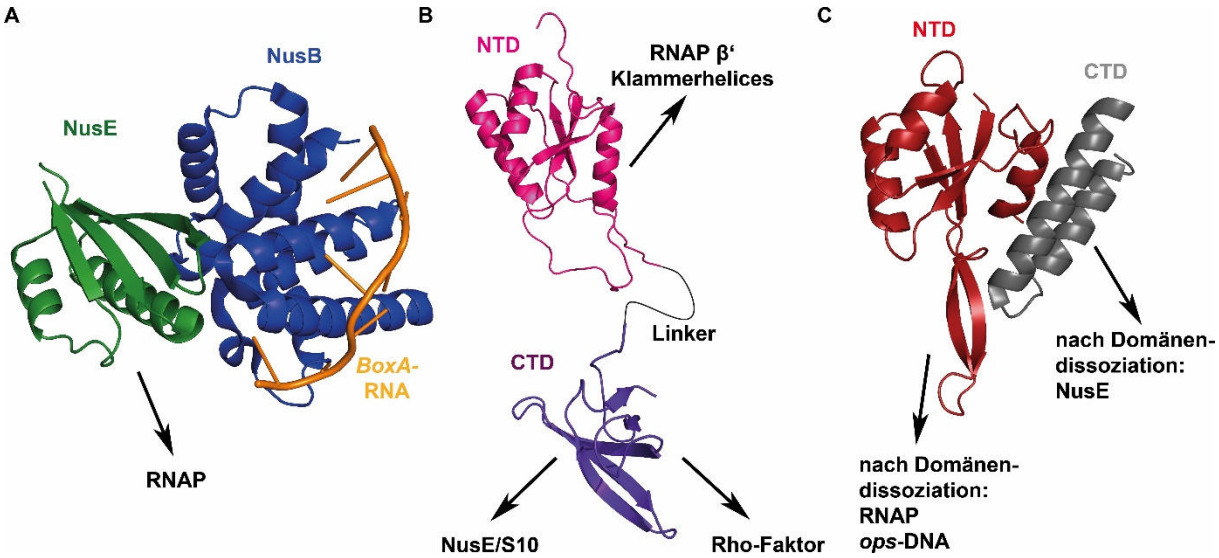


Abbildung 7: Strukturen von A) NusB:NusE mit gebundener *BoxA*-RNA, B) NusG und C) RfaH. Die RNA- und Proteinstrukturen sind in Cartoondarstellung abgebildet. Die Pfeile zeigen an mit welchem Partner bzw. mit welchen Partnern das jeweilige Protein oder die jeweilige Domäne interagieren kann. **A)** NusB (dunkelblau) bildet zusammen mit NusE (grün) ein Heterodimer, welches in der Lage ist *BoxA*-RNA (orange) zu binden (PDB-Code: 3R2C, Stagno *et al.*, 2011). **B)** In *E. coli* besteht NusG aus einer NTD und einer CTD, welche durch einen 15 Aminosäuren langen, flexiblen Linker verbunden sind (pink und lila, PDB-Codes: 2JVV, 2K06, Mooney *et al.*, 2009b). **C)** RfaH liegt in *E. coli* autoinhibiert vor, da die CTD (dunkelgrau) an die NTD (rot) gebunden ist (PDB-Code: 2OUG, Belogurov *et al.*, 2007). Die jeweiligen Interaktionen können erst nach Domänenöffnung erfolgen.

1.2.3 NusG und RfaH

Während die Kernstruktur der RNAP in Bakterien, Archaeen und Eukaryoten konserviert ist, trifft dies bei den Transkriptionsfaktoren nur auf NusG zu (Werner, 2012). *E. coli* NusG (*Ec*NusG) ist essentiell und besteht aus den beiden Domänen NTD und CTD, welche durch einen flexiblen Linker miteinander verknüpft sind (Abb. 7B, Downing *et al.*, 1990; Mooney *et al.*, 2009b; Burmann *et al.*, 2011). Die NTD bindet an den βGL und die β'CH der RNAP, wodurch die beiden Scheren der RNAP zusammengehalten werden (Martinez-Rucobo *et al.*, 2011; Sevostyanova *et al.*, 2011). Dies führt zu

einer stärkeren DNA-Bindung, zum Unterdrücken von Pausen und zur Erhöhung der Elongationsrate. *Ec*NusG-CTD interagiert mit S10 (NusE), welches Teil des Ribosoms ist und koppelt so Transkription und Translation (Burmann *et al.*, 2010). Des Weiteren rekrutiert die CTD den Transkriptionsterminationsfaktor Rho an die RNAP und ist notwendig für die Rho-abhängige Termination (Burmann *et al.*, 2010; Pasman und von Hippel, 2000; Sullivan und Gottesman, 1992). In *E. coli* gibt es zusätzlich den operonspezifischen Transkriptionsfaktor RfaH, welcher ein Paralog zu NusG ist (Bailey *et al.*, 1997). Bei diesem ist die Struktur der NTD ähnlich wie bei NusG. Die CTD ist jedoch fest an die NTD gebunden und besteht im Gegensatz zu NusG-CTD aus zwei α-Helices anstelle des β-Fasses (Abb. 7C, Belogurov *et al.*, 2007). Durch die Interaktion zwischen den beiden Domänen sind die Bindungsflächen für RNAP und andere Faktoren verdeckt. RfaH-CTD kann sich in eine β-Fassstruktur umwandeln, wenn sich die Domänen öffnen (Burmann *et al.*, 2012). Die Domäneninteraktion wird vermutlich durch Bindung der RfaH-NTD an die RNAP aufgehoben, welche an der *operon polarity suppressor* (*ops*-) DNA pausiert (Tomar *et al.*, 2013). Die *ops*-Sequenz befindet sich nahe des Promotors in der 5' nichttranslatierten Region von Operons, welche stark durch RfaH aktiviert werden. Die Aktivierung ist möglich, indem die Rho-abhängige Termination durch Ausschluss von NusG verhindert wird und gleichzeitig die Translation durch die Rekrutierung des Ribosoms aktiviert wird (Sevostyanova *et al.*, 2011; Belogurov *et al.*, 2009).

Strukturell gibt es bei NusG zwischen den Organismen einige Unterschiede. Ähnlich wie bei *Ec*NusG finden bei NusG aus *T. thermophilus* (*Tt*NusG) und *Aquifex aeolicus* (*Aa*NusG) keine intramolekularen Wechselwirkungen zwischen der NTD und CTD statt (Reay *et al.*, 2004; Steiner *et al.*, 2002). In NusG aus *T. maritima* (*Tm*NusG) interagieren jedoch die beiden Domänen miteinander in Lösung, wodurch die Thermostabilität erhöht wird (Drögemüller *et al.*, 2013). *Tm*NusG und *Aa*NusG haben beide außerdem eine Insertion von ca. 70 Aminosäuren in der NTD, welche eine eigene Domäne mit bisher unbekannter Funktion ausbilden. Für *Tm*NusG wird vermutet, dass diese Domäne Nukleinsäuren binden kann. Von *Mycobacterium tuberculosis* NusG (*Mt*NusG) ist bisher keine Struktur mit atomarer Auflösung vorhanden. Es konnten allerdings funktionelle Unterschiede zu *Ec*NusG festgestellt werden. So stimuliert *Mt*NusG im Gegensatz zu *Ec*NusG die intrinsische Termination und ist nicht in der Lage Rho zu binden (Czyz *et al.*, 2014; Kalyani *et al.*, 2014).

1.2.4 Die N-abhängige Antitermination

Bei der N-abhängigen Antitermination bildet sich ein stabiler Komplex, wodurch weit von der *nut*-Bindungsstelle entfernte, stromabwärts gelegene intrinsische und Rho-abhängige Terminatoren überlesen werden können. Eine wichtige Rolle spielt hierbei das λN-Protein. Bei ausreichendem Überschuss an λN kann die RNAP in eine terminationsresistente Form umgewandelt werden und *in vitro* über Terminatoren hinweg transkribieren, sogar in Abwesenheit der *nut*-Bindestelle (Rees *et al.*, 1996; Salstrom und Szybalski, 1978). Verantwortlich hierfür ist die λN-CTD (73-107), welche in der Nähe des aktiven Zentrums an die RNAP bindet (Mogridge *et al.*, 1998; Cheeran *et al.*, 2007). Für

eine effektive Antitermination sind jedoch die *nut*-Bindungsstelle, bestehend aus *BoxA* und *BoxB*, sowie die vier Proteine NusA, NusB, NusE und NusG nötig (Downing *et al.*, 1990; Friedman *et al.*, 1984). λN bindet mit seiner argininreichen NTD (1-22) an *BoxB* und mit der Region 34-47 an NusA-AR1, wodurch der Terminationsschleife-stabilisierende Effekt von NusA aufgehoben wird (Gusarov und Nudler, 2001; Mogridge *et al.*, 1998; Tan und Frankel, 1995). Beim derzeitigen Modell (Abb. 8) bildet die naszierende RNA eine Schleifenstruktur aus, bei der NusA-NTD und NusG-NTD an die RNAP gebunden sind und NusE indirekt über NusG assoziiert ist (Mooney *et al.*, 2009a; Mah *et al.*, 1999; Belogurov *et al.*, 2009; Whalen und Das, 1990). Ausgebildet wird die RNA-Schleife dadurch, dass der NusB:NusE Heterodimer *BoxA* und *BoxB* bindet (Stagno *et al.*, 2011). Für eine zusätzliche Stabilisierung des Transkriptionselongationskomplexes (TEC) sorgt die Interaktion von NusA-SKK mit der Sequenz zwischen *BoxA* und *BoxB* und die Wechselwirkung von NusG-CTD mit NusE (Burmann *et al.*, 2010; Prasch *et al.*, 2009). Die Bindung von NusG-CTD an NusE bewirkt, dass NusG-CTD nicht mehr mit Rho interagieren kann und die Rho-abhängige Termination inhibiert ist (Burmann *et al.*, 2010). Obwohl die einzelnen Wechselwirkungen im Antiterminationskomplex nicht sehr stark sind, formt sich durch die Vielzahl der Interaktionen ein stabiler Komplex (Greenblatt *et al.*, 1993).

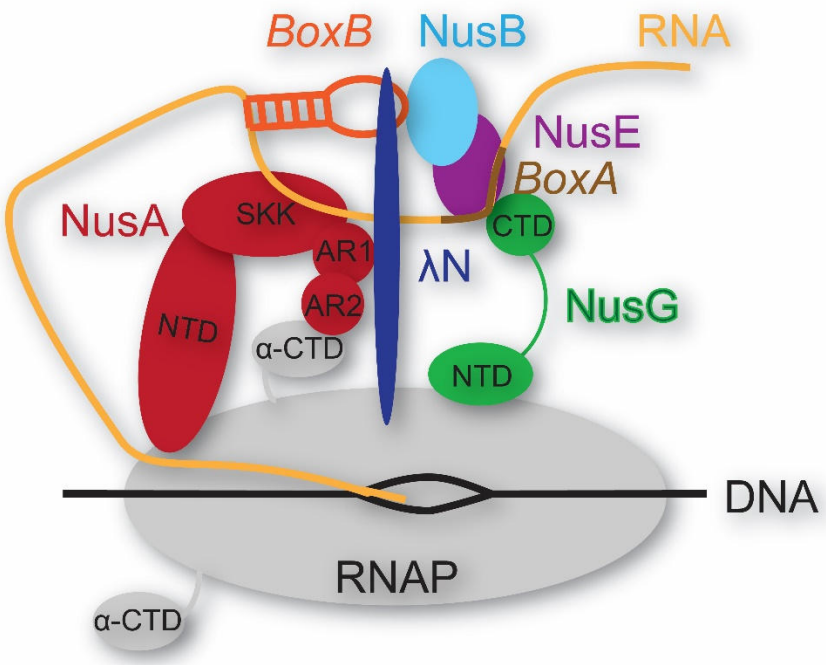


Abbildung 8: Schematische Darstellung der Interaktionen innerhalb des Antiterminationskomplexes.

1.3 Untersuchung großer Proteine und Proteinkomplexe mittels NMR-Spektroskopie

Grundlage der NMR-Spektroskopie ist, dass durch ein externes Magnetfeld die Energieniveaus zwischen den Orientierungen des magnetischen Dipols der Atomkerne aufgespalten werden (Zeeman-Effekt). Aufgrund dieser Aufspaltung ist das energieärmere Niveau stärker besetzt als das energiereichere Niveau. Dies führt zu einer makroskopischen Magnetisierung entlang des externen Magnetfeldes. Durch Radiofrequenzpulse können Übergänge zwischen den Energieniveaus erzeugt werden, sodass die Energieübergänge aller in der Probe vorhandenen Kerne gleichzeitig gemessen werden können. Die Resonanzfrequenzen der einzelnen Kerne unterscheiden sich, da die elektronische Umgebung und die Wechselwirkung zwischen benachbarten Atomkernen einen Einfluss auf die magnetische Umgebung des Spins ausüben. Diese Unterschiede in den Resonanzfrequenzen können zur Strukturaufklärung verwendet werden. Die Resonanzfrequenzen werden als chemische Verschiebung gegenüber einer Referenzsubstanz in *parts per million* (ppm) angegeben.

1.3.1 Strukturuntersuchungen kleinerer Proteine

Zur strukturellen Aufklärung von Proteinen sind eindimensionale (1D) Spektren aufgrund der vielen Signale und Überlagerungen nicht geeignet. Hierfür sind zwei- (2D) oder mehrdimensionale Spektren erforderlich. Für 2D Spektren wird die Magnetisierung vom Proton über skalare Kopplung (Bindung) auf Stickstoff oder Kohlenstoff übertragen. In den erhaltenen Spektren sind die chemischen Verschiebungen von zwei Kernen miteinander korreliert. Bei dreidimensionalen (3D) Spektren können mehrere chemische Verschiebungen korreliert werden, wodurch die Signale den einzelnen Aminosäuren zugeordnet werden können. Für die Strukturbestimmung werden zusätzlich Abstandsinformationen zwischen den Kernen benötigt. Hierzu wird ausgenutzt, dass der Magnetisierungstransfer nicht nur über skalare Kopplungen (Bindungen) sondern auch durch Dipol-Dipol-Wechselwirkungen oder Kreuzrelaxation über den Raum erfolgen kann (Kern-Overhauser-Effekt, NOE). Durch die Kombination einer Vielzahl solcher Abstandsinformationen lässt sich die dreidimensionale Struktur eines Proteins berechnen. Durch diese Techniken können Proteinstrukturen bis ca. 25 kDa gelöst werden (Abb. 9).

1.3.2 Methoden zur Untersuchung mittelgroßer Proteine

Mit zunehmender Proteingröße wird es schwieriger Proteinstrukturen mittels NMR-Spektroskopie zu bestimmen, da es aufgrund der Vielzahl an Resonanzen zu Signalüberlagerungen kommt und die Magnetisierung so schnell relaxiert, dass ein Großteil des Signals vor der Detektion verloren geht. Die Spektrenqualität ist daher für eine Zuordnung der Resonanzsignale nicht ausreichend, da zu wenige Abstandsinformationen erhalten werden. Durch Spektrometer mit höheren Feldstärken, den Einsatz von leistungsstarken Cryoprobenköpfen und neuen, verbesserten Pulstechniken, wie *Transverse relaxation optimized spectroscopy* (TROSY) können diese Probleme teilweise behoben werden. Des Weiteren kann durch Deuterieren des Proteins das Proton-Proton-Netzwerk ausgedünnt werden,

wodurch Dipol-Dipol-Wechselwirkungen verringert werden und die Magnetisierung langsamer relaxiert (Plevin und Boisbouvier, 2012). Außerdem kann die Komplexität der Spektren durch die selektive Markierung einzelner Aminosäuren verringert werden.

Mit zunehmender Proteingröße ist es allerdings auch mit diesen Techniken nicht mehr möglich die Struktur zu bestimmen. Das größte Protein bei dem auf diese Weise die Signale des Proteinrückgrats zugeordnet werden konnten, ist die Malatsynthase G (81,4 kDa, Abb. 9, Tugarinov *et al.*, 2002). Bei größeren Systemen war dies nur möglich, wenn das Zielprotein als Homooligomer vorliegt (Salzmann *et al.*, 2000; Fiaux *et al.*, 2002).

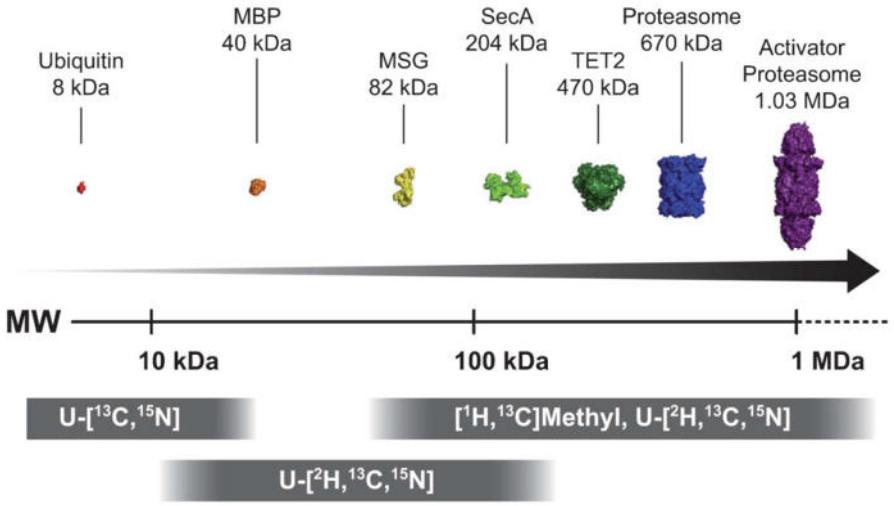


Abbildung 9: Größe der durch NMR-Spektroskopie untersuchten Proteine und Proteinkomplexe. Die Proteine sind gegen ihr Molekulargewicht aufgetragen. Folgende PDB-Codes wurden verwendet: Ubiquitin, 1UBQ; maltosebindendes Protein, (MBP), 1DMB; Malatsynthase G (MSG), 1P7T; SecA, 2VDA; TET2, 2WZN; Proteasom, 3OKJ; Proteasom–Aktivator-Komplex, 1Z7Q. Die für die Untersuchung geeigneten Markierungsstrategien sind unten aufgeführt, wobei U für das gesamte Protein bzw. den Rest des Proteins steht. Abbildung entnommen aus Plevin und Boisbouvier, 2012

.

1.3.3 Untersuchung großer Proteine bzw. Proteinkomplexe mittels NMR

Bei großen Proteinen und Proteinkomplexen ist es, aufgrund der schnellen Magnetisierungsrelaxation und der Vielzahl an Signalüberlagerungen, nicht mehr möglich ausreichend viele Strukturinformationen für eine *de novo* Strukturbestimmung zu gewinnen. Die NMR-Spektroskopie ist hier daher besonders von Bedeutung, um Aussagen über Funktion, Dynamik und Interaktionspartner zu erhalten. Der Grund für die schnelle Relaxation der Magnetisierung ist, dass durch die Molekülgröße die Rotation verlangsamt ist und der Magnetisierungstransfer effizienter abläuft. Dadurch geht ein Großteil der Magnetisierung vor der Detektion verloren und es kommt zu einer starken Verbreiterung bzw. zum Verschwinden der NMR-Signale.

Um große Systeme zu untersuchen, können Methylgruppen bestimmter Aminosäurereste spezifisch ^{1}H, ^{13}C markiert werden, während der Rest des Proteins deuteriert und ansonsten unmarkiert vorliegt (Methylgruppenmarkierung, Ruschak und Kay, 2010). Methylgruppen sind besonders sensitiv, da

immer drei Methylprotonen vorliegen, deren Signale aufgrund der Drehung um die Methylsymmetrieachse entartet sind (Nicholson *et al.*, 1992). Außerdem befinden sie sich oft am Ende langer Aminosäureseitenketten und sind daher dynamischer als Rückgratamide, wodurch die Magnetisierung langsamer relaxiert. Des Weiteren sind sie über die gesamte Proteinsequenz verteilt und befinden sich sowohl im hydrophoben Inneren des Proteins als auch an Interaktionsflächen am Äußeren des Proteins (Plevin und Boisbouvier, 2012; Rosen *et al.*, 1996; Gardner *et al.*, 1997). Für die Markierung werden entsprechend markierte Aminosäuren oder Aminosäurevorstufen den Medien kurz vor der Induktion zugesetzt. Die Zugabe von α-Ketobutyrat führt beispielsweise zur Markierung von Ile-Resten und die Zugabe von α-Ketoisovalerat zur Markierung von Val- und Leu-Resten (Abb. 10).

Mit den erhaltenen Spektren können, ebenso wie bei anderen NMR-Experimenten, Abstands-, Struktur-, Funktions- und Dynamikmessungen durchgeführt werden. Voraussetzung hierfür ist eine sequenzspezifische Zuordnung der Resonanzsignale. Um dies zu erreichen kann das System in kleinere, isolierte Bausteine zerlegt werden und die Resonanzsignale zugeordnet werden. Anschließend wird der Komplex zusammengebaut und die Zuordnung angepasst. Hierdurch konnten nahezu alle Resonanzen der α-Untereinheit (21 kDa) im 20S-Proteasom (670 kDa) zugeordnet werden und intermolekulare Bindungsflächen und Bindungsflächen zum 11S-Aktivatorkomplex nachgewiesen werden (Sprangers und Kay, 2007). Die isolierten Untereinheiten sollten sich hierbei strukturell nicht stark von den im Komplex Vorkommenden unterscheiden, da sich die Resonanzsignale ansonsten stark verändern. Eine weitere Möglichkeit die Resonanzen zuzuordnen besteht darin die isotopenmarkierten Aminosäuren einzeln zu einer nicht markierten auszutauschen. Beim entsprechenden Spektrum sollte dann nur das Signal dieser Aminosäure fehlen. Diese Technik wurde beim homododecameren archaelen Protein TET2 eingesetzt (Amero *et al.*, 2011). Beide Zuordnungstechniken können auch kombiniert werden und es ist von Vorteil bereits eine Kristallstruktur als Grundlage zu haben. Der Schwerpunkt der NMR-Spektroskopie liegt daher bei großen Proteinkomplexen nicht zwingend in der Strukturaufklärung, sondern in der Untersuchung von niedrig besetzten Zuständen, Interaktionen oder Domänenbewegungen.

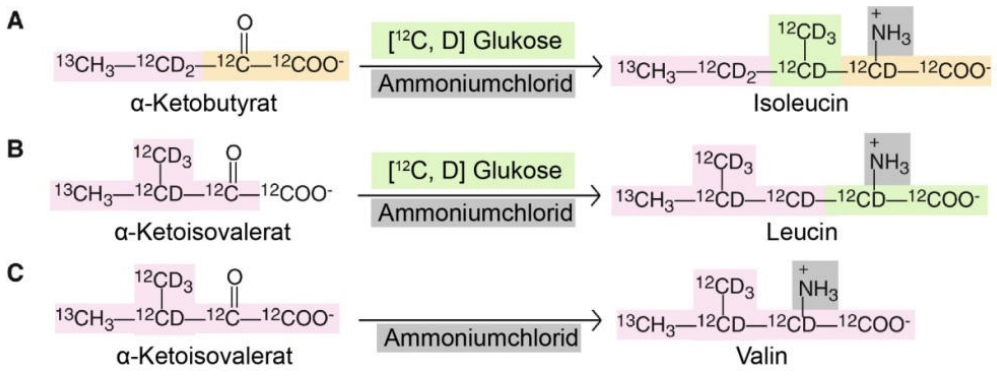


Abbildung 10: Aminosäurevorstufen, welche zu Minimalmedium zugegeben werden müssen, um bestimmte Methylgruppen von A) Isoleucin, B) Leucin und C) Valin ^{13}C, ^{1}H zu markieren, während der Rest des Proteins ^{12}C, ^{2}H markiert ist. Die farblichen Hinterlegungen zeigen an, wie die Vorstufen, Ammoniumchlorid und Glukose in die jeweilige Aminosäure eingebaut werden. Die Abbildung wurde, leicht modifiziert, aus Ruschak und Kay, 2010 entnommen.

2. Zielsetzung

Die Transkription ist einer der zentralen Prozesse in der Zelle, wobei die RNAP das wichtigste Enzym ist. Ihre Regulation, welche eine hohe Flexibilität bei der Genexpression ermöglicht, ist jedoch weitgehend unbekannt. In Bakterien spielen die Nus-Faktoren hierbei eine entscheidende Rolle. Ziel dieser Arbeit war es daher, Wechselwirkungen der Nus-Faktoren mit der RNAP und der Nus-Faktoren untereinander auf molekularer Ebene zu untersuchen.

NusG nimmt bei der Rho-abhängigen Termination eine zentrale Rolle ein und verknüpft die Transkription mit der Translation. Da *Mt*NusG andere Effekte auf die Transkription ausübt als *Ec*NusG, sollte die *Mt*NusG-Struktur gelöst werden, um zu untersuchen ob strukturelle Unterschiede hierfür verantwortlich sind (Czyz *et al.*, 2014). Da die Domänen in NusG je nach Organismus fest aneinander binden oder flexibel verknüpft sind, sollte zusätzlich die Domänenwechselwirkung in *Mt*NusG analysiert werden.

Für die Regulation der Transkription sind häufig direkte Interaktionen mit der RNAP verantwortlich, aber auch Wechselwirkungen der Faktoren untereinander können die Transkription beeinflussen. In dieser Arbeit sollte daher geprüft werden, ob die beiden Transkriptionsfaktoren NusA und NusG aus *E. coli* miteinander interagieren und welche Auswirkungen eine solche Interaktion auf die Elongation oder Termination haben kann.

Da viele Wechselwirkungen direkt mit der RNAP stattfinden, sollte ein System für *E. coli* entwickelt werden, um zu ermitteln, an welche Untereinheit der RNAP ein Interaktionspartner bindet. Hierzu sollten die Gene der einzelnen RNAP-Untereinheiten getrennt exprimiert und die entsprechenden Proteine gereinigt werden. Durch Titrationen der ^{15}N-markierten Nus-Faktoren (NusA-NTD, NusA-AR2, NusG-NTD, NusE) mit den isolierten Untereinheiten sollten prognostizierte Bindungsstellen verifiziert und neue bestimmt werden.

Ein weiteres Ziel war es die Grundlage zu schaffen, um Dynamiken und schwache Wechselwirkungen der RNAP zu untersuchen. Hierzu sollten NMR-Spektren der RNAP bzw. einzelner Untereinheiten in der RNAP aufgenommen werden, indem einzelne RNAP-Untereinheiten methylgruppenmarkiert werden und ein Protokoll für die funktionelle Assemblierung der RNAP *in vitro* entwickelt wird.

Die Bindungsflächen an die RNAP auf Seiten der Nus-Faktoren erlauben eine Aussage über die Orientierung am Transkriptionskomplex. Außerdem zeigen sie, welche Interaktionen gleichzeitig oder exklusiv erfolgen können und lassen Rückschlüsse auf die Mechanismen und Funktionsweisen der Nus-Faktoren zu. Jedoch sind die Interaktionsflächen größtenteils unbekannt. Durch Methylgruppen-markieren von NusA, NusE und NusG und anschließender Titration mit der RNAP sollte eine Methode entwickelt werden, um die Bindungsflächen an die RNAP zu bestimmen.

3. Zusammenfassung und Diskussion der Ergebnisse

3.1 Strukturbestimmung von NusG aus *Mycobacterium tuberculosis*

Als einziger Transkriptionsfaktor kommt NusG in Bakterien und Archaeen vor und ist als Homolog Spt5 in Eukaryoten vorhanden (Werner, 2012). *Ec*NusG-NTD interagiert mit der RNAP und erhöht die Transkriptionsrate, während *Ec*NusG-CTD an NusE/S10 oder an Rho binden kann (Mooney *et al.*, 2009b; Burmann *et al.*, 2010). Die Strukturen von NusG aus *E. coli*, *T. thermophilus*, *T. maritima* und *Aquifex aeolicus* waren bereits bekannt. Bei *Ec*NusG, *Tt*NusG und *Aa*NusG sind keine intramolekularen Domänenwechselwirkungen vorhanden (Burmann *et al.*, 2011; Reay *et al.*, 2004; Steiner *et al.*, 2002). Bei *Tm*NusG ist jedoch die NTD fest mit der CTD verbunden, wodurch es in einem autoinhibierten Zustand vorliegt, da die Bindestellen an Rho, NusE und RNAP verdeckt sind (Drögemüller *et al.*, 2013). Funktionell unterscheiden sich *Mt*NusG und *Ec*NusG. So stimuliert *Mt*NusG die intrinsische Termination, während *Ec*NusG auf diese keinen Einfluss hat (Czyz *et al.*, 2014). Außerdem bindet *Mt*NusG nicht an Rho, jedoch an NusE (Kalyani *et al.*, 2014). Um zu untersuchen ob hierfür strukturelle Unterschiede verantwortlich sind, wurde die Struktur von *Mt*NusG in Einzelarbeit A analysiert.

Im Vergleich zu anderen NusG-Proteinen befinden sich bei *Mt*NusG am N-Terminus 40 zusätzliche Aminosäuren und der Linker zwischen NTD und CTD ist verlängert. Erste [^{1}H, ^{15}N] *Heteronuclear Single Quantum Coherence* (HSQC-) Spektren von *Mt*NusG zeigten ein für gefaltete Proteine typisches Spektrum, in welchem die Resonanzfrequenzen breit verteilt sind (Abb. 11). Zusätzlich sind jedoch mehrere scharfe, intensive Peaks bei chemischen Verschiebungen vorhanden, die typisch für Aminosäuren in unstrukturierten Bereichen sind. Diese Signale gehören vermutlich zu den zusätzlichen Aminosäuren am N-Terminus und der Linkerregion. Die daraus resultierenden Signalüberlagerungen und die erhöhte Rotationskorrelationszeit ließen keine Strukturbestimmung des Gesamtproteins zu. Daher wurden die beiden Domänen getrennt produziert und analysiert (*Mt*NusG-NTD: Aminosäuren 1-178, *Mt*NusG-CTD: Aminosäuren 178-238) und anschließend mit dem Gesamtprotein verglichen.

Das [^{1}H, ^{15}N]-HSQC-Spektrum von *Mt*NusG-NTD zeigt ebenfalls die für ein gefaltetes Protein typischen chemischen Verschiebungen (Abb. 11A). Da das Protein jedoch über einen längeren Zeitraum instabil war, wurden für die Strukturvorhersage *in silico*-Methoden verwendet. Programme zur Sekundärstrukturvorhersage wie PSIPRED (Jones, 1999) prognostizierten, dass 34 Aminosäuren am N-Terminus unstrukturiert vorlägen. Dies ist in guter Übereinstimmung mit dem gemessenen [^{1}H, ^{15}N]-HSQC-Spektrum des Volllängenproteins. Die Programme zur Tertiärstrukturvorhersage I-TASSER (Roy *et al.*, 2010) und PHYRE2 (Kelley und Sternberg, 2009) kamen zu dem einheitlichen Ergebnis, dass der N-Terminus unstrukturiert sei und der restliche Teil der *Mt*NusG-NTD den bekannten NusG-NTD-Strukturen ähnele (Abb. 12A, B). Für den N-Terminus wurde vermutet, dass er

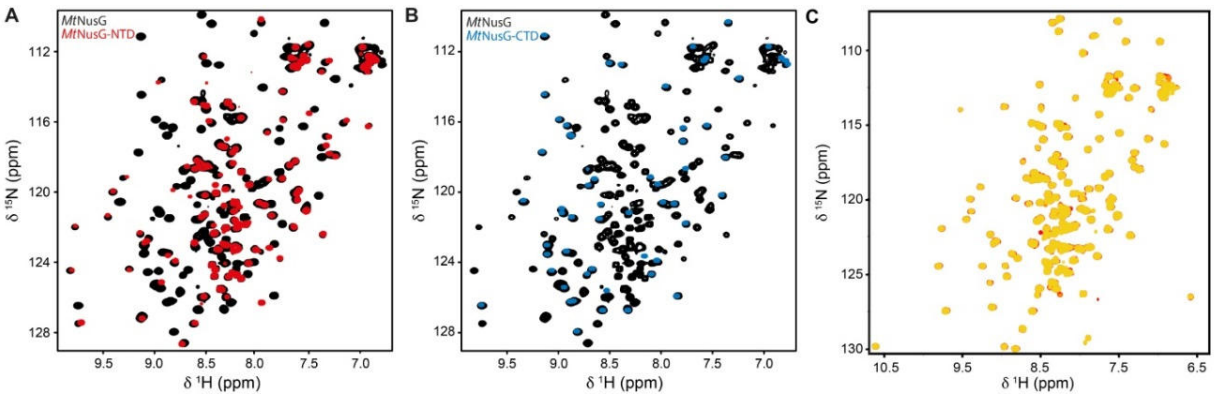


Abbildung 11: HSQC-Spektren der Konstrukte *Mt*NusG, *Mt*NusG-NTD und *Mt*NusG-CTD. A) Überlagerung der [^{1}H, ^{15}N]-HSQC-Spektren von ^{15}N-*Mt*NusG (schwarz) und ^{15}N-*Mt*NusG-NTD (rot). **B)** Überlagerung der [^{1}H, ^{15}N]-HSQC-Spektren von ^{15}N-*Mt*NusG (schwarz) und ^{15}N-*Mt*NusG-CTD (blau). **C)** [^{1}H, ^{15}N]-HSQC-Spektrum von ^{15}N-*Mt*NusG-NTD in Abwesenheit (rot) und Anwesenheit von *Mt*NusG-CTD (1:1, orange; 1:2 gelb)

hydrophobe Bereiche an der NTD und CTD verdeckt, um die Löslichkeit des Proteins zu erhöhen (Kalyani *et al.*, 2014). Dies ist jedoch unwahrscheinlich, da sich die elektrostatischen Eigenschaften von *Mt*NusG und *Ec*NusG stark ähneln und eine Maskierung der hydrophoben Bereiche daher nicht notwendig erscheint (Einzelarbeit A, Abb. S3). Der Aminoterminus besteht im Wesentlichen aus Alaninen und polaren, meist sauren, Resten (12 Ala, 4 Glu, 5 Asp, 4 Thr). Es könnte sich daher um einen natürlichen Löslichkeitsanhang handeln. Gestützt wird diese Vermutung dadurch, dass Versuche durch mich und eine andere Arbeitsgruppe ein aminoterminal verkürztes Konstrukt zu generieren, zur Präzipitation des Protein führten (Kalyani *et al.*, 2014).

Für die Strukturaufklärung der *Mt*NusG-CTD wurden mit einer ^{1}H, ^{13}C, ^{15}N-markierten Proteinprobe 3D NMR-Standardexperimente gemessen, die Resonanzen zugeordnet und, unter Zuhilfenahme von NOEs, die Struktur berechnet (Abb. 12C, D). Sie besteht aus fünf β-Strängen (β1: Ser190-Val193; β2: Pro202-Asn209; β3: Lys214-Val219; β4: Thr226-Thr231; β5: Val235-Ile238), welche eine antiparallele, fassartige β-Faltblattstruktur ausbilden. Ein Vergleich der *Mt*NusG-CTD mit den entsprechenden Abschnitten der Strukturen von *Ec*NusG (PDB-Code: 2JVV), *Aa*NusG (PDB-Code: 1M1G), *Tt*NusG (PDB-Code: 1NZ9) und *Tm*NusG (PDB-Code: 2LQ8) ergab *root mean square deviation* (rmsd-) Werte der Rückgratatome von 0,9-1,4 Å, was einer hohen strukturellen Ähnlichkeit entspricht. Für *Mt*NusG-CTD wurde aufgrund von Circulardichroismus- (CD-) Daten eine für NusG-CTDs ungewöhnliche, verzerrte Struktur prognostiziert (Kalyani *et al.*, 2014). Dies konnte jedoch durch die Strukturbestimmung nicht bestätigt werden.

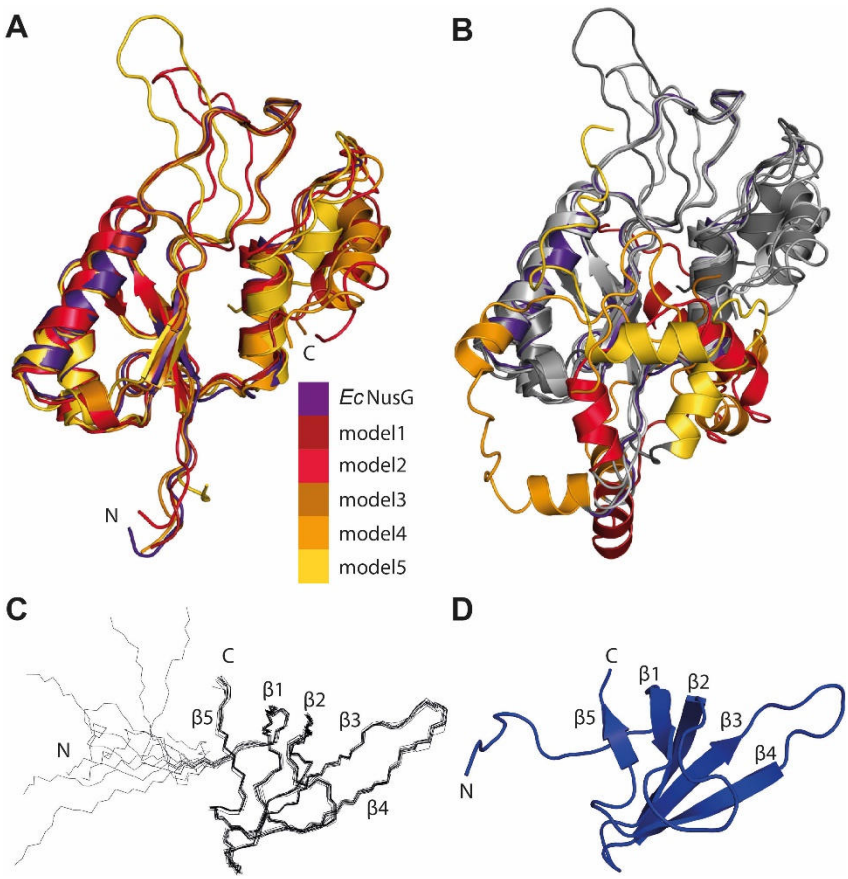


Abbildung 12: Strukturen der *Mt*NusG-NTD (A, B) und *Mt*NusG-CTD (C, D). A) Cartoondarstellung der fünf mittels I-TASSER berechneten Homologiemodelle für die *Mt*NusG-NTD (dunkelrot, pink, braun, orange, gelb) und die Struktur der *Ec*NusG-NTD (PDB-Code: 2K06, lila, Mooney *et al.*, 2009b; Roy *et al.*, 2010). Zur besseren Übersicht sind die ersten 40 Aminosäuren nicht gezeigt. **B)** In grau sind die *Mt*NusG-NTD Strukturen aus A) gezeigt. Zusätzlich sind die 40 aminoterminalen Aminosäuren in der Farbe des jeweiligen Modells eingezeichnet. **C)** Bänderdarstellung des Proteinrückgrats der zehn energetisch besten *Mt*NusG-CTD-Strukturen. Die Strukturen wurden mittels NMR-Spektroskopie bestimmt. Der unstrukturierte Bereich entspricht dem aminoterminalen Ende der *Mt*NusG-CTD und ist im Gesamtprotein der Linker zwischen NTD und CTD. **D)** Cartoondarstellung der energetisch günstigsten Struktur. Die Termini und die Sekundärstrukturelemente sind angegeben.

Zur Überprüfung, ob die beiden *Mt*NusG-Domänen miteinander interagieren, wurden zunächst die aufgenommenen [^{1}H, ^{15}N]-HSQC-Spektren des Volllängenproteins mit denen der isolierten Domänen verglichen (Abb. 11). Die chemischen Verschiebungen der *Mt*NusG-CTD und der *Mt*NusG-NTD sind mit denen des Gesamtproteins nahezu identisch. Daher kann davon ausgegangen werden, dass die beiden Domänen nicht, oder zumindest nicht stark, interagieren. Zum Nachweis von möglichen transienten Domänenwechselwirkungen, wurde ^{15}N-markierte *Mt*NusG-NTD mit *Mt*NusG-CTD titriert. Hierbei wurden keine Hinweise auf eine Interaktion beobachtet. Es kam weder zu einer signifikanten Signalabnahme, noch zu Veränderungen der chemischen Verschiebungen (Abb. 11C). Die Unabhängigkeit der beiden Domänen wurde zusätzlich durch NMR-Relaxationsmessungen mit dem Gesamtprotein untersucht. Die Ergebnisse zeigen eine trimodale Verteilung (Einzelarbeit A, Abb. 5). Diese können der *Mt*NusG-CTD, der *Mt*NusG-NTD und weiteren flexiblen Bereichen zugordnet

werden. Hierdurch wird bestätigt, dass flexible Bereiche vorhanden sind, die Domänen unabhängig voneinander rotieren und daher nicht miteinander wechselwirken.

Da bisher nur bei *Tm*NusG eine Domänenwechselwirkung und die Maskierung der Bindestellen von NusE, Rho und RNAP bekannt sind, handelt es sich hierbei vermutlich um kein generelles Konzept. Bei *Tm*NusG trägt die Domäneninteraktion zur Stabilisierung des Gesamtproteins bei (Drögemüller *et al.*, 2013). Diese wird benötigt, da *T. maritima* ein Temperaturoptimum von 80 °C besitzt (Huber *et al.*, 1986). Der terminationsverstärkende Effekt von *Mt*NusG ist weder auf den verlängerten Linker, noch auf den zusätzlichen Bereich am Aminoterminus zurückzuführen (Czyz *et al.*, 2014). Da, wie von mir gezeigt, die elektrostatischen Eigenschaften und die Struktur von *Mt*NusG denen von *Ec*NusG ähneln, gehe ich davon aus, dass die funktionalen Effekte auf einer unterschiedlichen Bindung an die RNAP im Vergleich zu *Ec*NusG beruhen. Basierend auf den variierenden Effekten von NusG in unterschiedlichen Organismen, wurde ein Modell entwickelt, wie NusG-CTD unterschiedliche RNAP modulierende Faktoren binden kann und diese über die NusG-NTD an die RNAP bringen kann. Die spezifischen Funktionen müssen jedoch funktionell und strukturell bestimmt werden (Sevostyanova und Artsimovitch, 2010).

3.2 Interaktion der Transkriptionsfaktoren NusA und NusG

Die beiden Transkriptionsfaktoren NusA und NusG spielen bei der bakteriellen Transkription eine wichtige Rolle und können entweder synergistisch oder antagonistisch arbeiten. Beide Faktoren sind Teil des Antiterminationskomplexes, regulieren die Synthesegeschwindigkeit der RNAP, das Pausieren der RNAP und beeinflussen sowohl die intrinsische als auch die Rho-abhängige Termination (zusammengefasst in Roberts *et al.*, 2008). An beiden Transkriptionsfaktoren wird seit mehreren Jahrzehnten intensiv geforscht, wobei die meisten der Untersuchungen in *E. coli* durchgeführt wurden. Im Folgenden sind daher, wenn von den Nus-Faktoren oder der RNAP geschrieben wird, soweit nicht anders vermerkt, die jeweiligen *E. coli* Proteine gemeint. Bisher war keine Wechselwirkung der Proteine NusA und NusG untereinander bekannt und oftmals wurde nur der Einfluss einer der beiden Transkriptionsfaktoren untersucht. Lediglich für zwei Terminatoren wurde gezeigt, dass NusG und NusA unabhängig voneinander an die RNAP binden und ihren Effekt ausüben können (Burns *et al.*, 1998). Bei Terminationsassays an Rho-abhängigen Terminatoren hingegen wurde ein synergistischer Effekt beobachtet (Muteeb *et al.*, 2012). In diesem Fall hatte die Zugabe von NusG oder NusA alleine kaum einen Einfluss auf das Terminationsverhalten der RNAP, während sich das Terminationsmuster in Anwesenheit beider Faktoren dramatisch änderte.

In Einzelarbeit B wurde untersucht, ob NusG und NusA aus *E. coli* interagieren. Hierfür wurde zunächst ein [^{1}H, ^{15}N]-HSQC von ^{15}N-NusG vor und nach Zugabe von NusA gemessen. Während die Signalintensitäten der NusG-NTD in Gegenwart von NusA stark abnahmen, waren die Signale der NusG-CTD nahezu unbeeinflusst. Dies deutet auf eine Wechselwirkung zwischen NusG-NTD und NusA hin, da die Größenzunahme durch die NusA-Bindung zu einer schnelleren Relaxation der

Magnetisierung und damit zu einem Signalverlust führt. Die NusG-CTD ist im Gegensatz dazu weiterhin flexibel. Durch Titrationen der isolierten, ^{15}N-markierten NusG-Domänen mit NusA wurde die Interaktion bestätigt. Erneut war bei der NusG-NTD nach NusA-Zugabe eine starke Signalabnahme zu beobachten, während die Signale der NusG-CTD von NusA kaum beeinflusst wurden. Die interagierende Domäne bei NusA wurde identifiziert, indem zu ^{15}N-NusG-NTD ein zweifacher Überschuss der einzelnen NusA-Domänen (NTD, SKK, AR1 und AR2) zugegeben wurde. Lediglich bei NusA-AR2-Zugabe kam es zu Veränderungen der chemischen Verschiebungen. Dies deutet auf eine Interaktion von NusA-AR2 und NusG-NTD hin, was durch einen *Pulldown* bestätigt wurde.

Das Vorgehen zur Ermittlung der Bindungsflächen und des K_D-Wertes mittels NMR-Spektroskopie ist im Folgenden für NusG-NTD gezeigt. Die Messungen mit NusA-AR2 erfolgten analog (siehe Einzelarbeit B). Für die Bestimmungen wurde ^{15}N-NusG-NTD vorgelegt und mit NusA-AR2 titriert, wobei nach jedem Titrationsschritt ein [^{1}H, ^{15}N]-HSQC-Spektrum aufgenommen wurde (Abb. 13A, B). Hierbei kommt es zu signifikanten Verschiebungen der Resonanzen einzelner Reste. Durch die Frequenzänderungen ($\Delta\nu$) konnte, unter Annahme eines Zweizustandsmodells, ein K_D-Wert von 17 µM für die Bindung ermittelt werden (Abb. 13C). Für die einzelnen Aminosäuren wurden außerdem, durch Vergleich der chemischen Verschiebungen des Anfangs- und Endzustandes der Titration, die normierten Änderungen der chemischen Verschiebungen ($\Delta\delta_{norm}$) errechnet (Abb. 13D). Für die Bestimmung der Bindungsflächen wurden Reste, bei welchen $\Delta\delta_{norm}$ einen bestimmten Schwellenwert überschritten hatte, auf der NusG-NTD-Struktur markiert (Abb. 13E). Durch die Software HADDOCK (de Vries *et al.*, 2010) wurde schließlich ein Modell für die Bindung von NusG-NTD an NusA-AR2 erhalten (Abb. 13F). Bei NusA-AR2 sind vor alle die C-terminalen Reste Trp 490 und Phe 491 betroffen, während bei NusG-NTD Signale der C-terminalen Helix und in der verlängerten Schleifenregion betroffen waren.

Während NusG-NTD an die β‘CH und die βGL der RNAP bindet, interagiert NusA-AR2 mit der RNAP αCTD (Mah *et al.*, 2000; Martinez-Rucobo *et al.*, 2011; Sevostyanova *et al.*, 2011). Daher wurde überprüft, ob die Interaktion zwischen NusA-AR2 und NusG-NTD auch in Gegenwart der RNAP stattfindet. Durch Verdrängungsexperimente konnte gezeigt werden, dass sowohl NusA-AR2 von der RNAP α-CTD und der gesamten RNAP, als auch die NusG-NTD von der gesamten RNAP durch den jeweils anderen Bindungspartner verdrängt werden kann (Einzelarbeit B, Abb. 2).

Die physiologische Bedeutung dieser Wechselwirkung der beiden Domänen wurde mittels eines *in vitro*-Transkriptionsassays untersucht (Einzelarbeit B, Abb. 3). Hierbei wurde die Verlängerung eines radioaktiv markierten RNA-Stückes während der Transkription in Abwesenheit oder in Gegenwart der verschiedenen Proteine bzw. Proteindomänen beobachtet. Durch NusG-NTD kam es erwartungsgemäß zu einer Unterdrückung von Transkriptionspausen, während durch NusA-AR2 eine Pause zu Beginn der Transkription verstärkt wurde. Bei gleichzeitiger Zugabe von NusA-AR2 und NusG-NTD kam es zur Ausbildung einer neuen Pause. Diese ist in Gegenwart von NusG und NusA-

AR2 bzw. beider Volllängenproteine deutlicher. Ein synergistischer Effekt von NusA und NusG wurde bereits für die Rho-abhängige Termination gezeigt (Muteeb *et al.*, 2012). Hierbei hatte die Zugabe von NusA oder NusG keinen Effekt auf die Termination, während die Gegenwart beider Proteine zu einer verfrühten Termination führte.

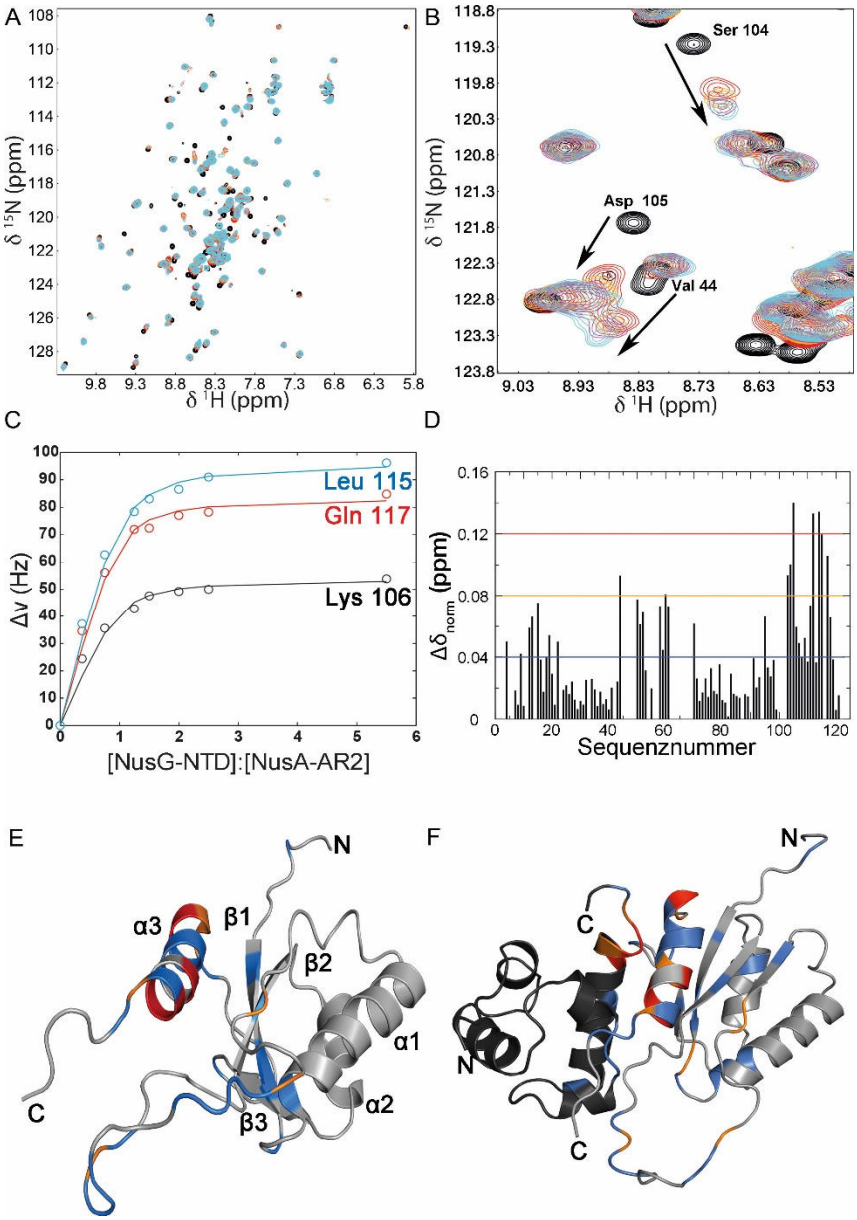


Abbildung 13: Bestimmung der Interaktionsflächen zwischen NusG-NTD und NusA-AR2. **A)** [^{1}H, ^{15}N]-HSQC-Titration von ^{15}N-NusG-NTD mit NusA-AR2. Das Spektrum von 140 µM ^{15}N-NusG-NTD ist in schwarz gezeigt. NusA-AR2 wurde in den molaren Verhältnissen 1:0,5 (rot), 1:1 (orange), 1:2 (pink) und 1:3 (hellblau) zugegeben. **B)** Vergrößerter Bereich aus A). Die Änderungen der chemischen Verschiebungen einiger Reste durch die Titration sind durch Pfeile angedeutet und die Signale sind beschriftet. **C)** Auftragung der Frequenzänderungen der Reste Gln 117, Lys 106 und Leu 115 in Abhängigkeit vom molaren Verhältnis von NusG-NTD und NusA-AR2. Die Linien entsprechen dem besten Fit unter Annahme eines Zweizustandsmodells. **D)** Auftragung der normierten Änderungen der chemischen Verschiebung ($\Delta\delta_{norm}$) von ^{15}N-NusG-NTD bei der Titration mit NusA-AR2 gegen die Sequenznummer. Die Signifikanzgrenzen für stark, mittel und schwach betroffene Reste sind: $\Delta\delta_{norm}$ (ppm) = 0,12: rot; $\Delta\delta_{norm}$ = 0,08: orange; $\Delta\delta_{norm}$ = 0,04: blau. **E)** Bindungsfläche von NusA-AR2 an NusG-NTD. Bei der Titration betroffene Regionen sind entsprechend der $\Delta\delta_{norm}$ markiert (siehe D). Die Struktur von NusG-NTD ist in Cartoondarstellung (hellgrau) gezeigt (PDB-Code: 2K06, Mooney *et al.*, 2009b). Die Sekundärstrukturelemente sind beschriftet. **F)** Modell des NusG-NTD:NusA-AR2-Komplexes. Nicht betroffene Reste der NusA-AR2 sind in dunkelgrau gezeigt (PDB-Code: 1WCN, Eisenmann *et al.*, 2005). Betroffene Reste der Titration von ^{15}N-NusA-AR2 mit NusG-NTD sind farblich markiert. Die Signifikanzniveaus entsprechen denen der Titration von ^{15}N NusG-NTD mit NusA-AR2.

Die Bildung des NusA:NusG-Komplexes könnte mehrere Effekte der Transkriptionsregulation erklären (Abb. 14). Zunächst ist NusA möglicherweise bei der Rekrutierung von freiem oder an das Ribosom gebundenem NusG an den TEC behilflich. ChIP-chip Daten zeigen, dass NusG und NusA erst mit der RNAP interagieren, nachdem der TEC die Promotorregion verlassen hat und der σ-Faktor dissoziiert ist (Mooney *et al.*, 2009a). NusG bindet hierbei später als NusA. NusA-AR2 würde somit dabei helfen NusG zu rekrutieren, indem durch die Interaktion die lokale Konzentration erhöht wird. Hierbei bleibt NusA über die NTD an die RNAP gebunden.

Eine weitere wichtige Aufgabe des Komplexes könnte die RNAP-Ribosom-Synchronisation zwischen Ribosom und RNAP sein. NusG verknüpft die beiden Prozesse, indem NusG-NTD mit der RNAP und NusG-CTD mit dem ribosomalen Protein S10 interagiert (Mooney *et al.*, 2009b; Burmann *et al.*, 2010). Besonders wenn die Synchronisation beider Prozesse, beispielsweise durch Pausen, gestört ist, könnte NusG aufgrund seiner geringen Größe für die Kopplung der beiden Prozesse nicht ausreichen. Ein vorübergehendes Ablösen der NusG-NTD vom TEC durch NusA-AR2 würde zu einem längeren RNAP-Ribosom-Verbindungsglied führen. Eventuell ist auch die Rekrutierung des Ribosoms zu Beginn der Transkription durch den verlängerten Linker vereinfacht.

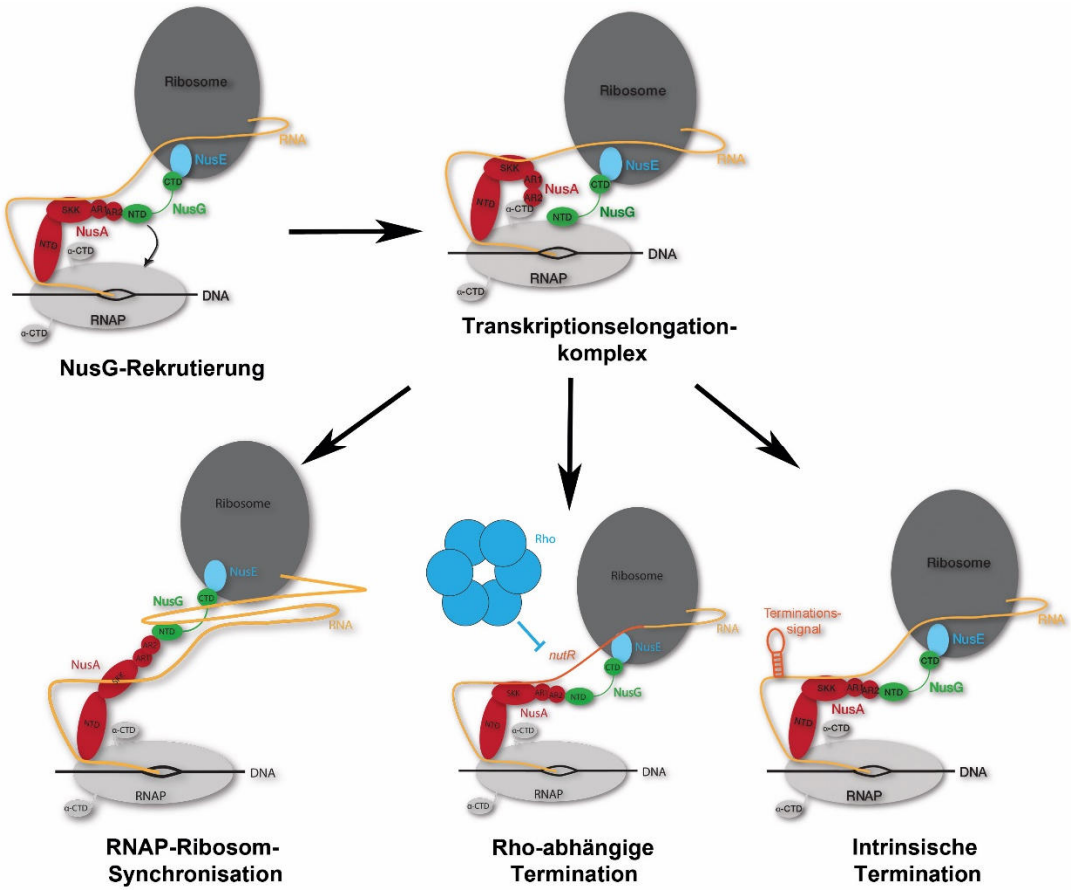


Abbildung 14: Mögliche Funktionen der NusG:NusA-Interaktion bei der Transkriptionselongation und –termination. Durch die NusG:NusA-Interaktion könnte NusG an den TEC rekrutiert werden. Außerdem könnte es zur Ausbildung einer längeren Verknüpfung zwischen Transkription und Translation kommen, beispielsweise wenn das Ribosom verlangsamt ist und NusG als Linker zu kurz wäre. Des Weiteren spielt die Interaktion zwischen NusA und NusG möglicherweise bei der Rho-abhängigen und intrinsischen Termination eine Rolle.

Die Interaktion zwischen NusG und NusA hat eventuell auch einen Einfluss auf die intrinsische oder Rho-abhängige Termination. Bei der intrinsischen Termination pausiert der TEC und es kommt anschließend zur Ausbildung der Terminationshaarnadelschleife (Weixlbaumer *et al.*, 2013). Der von uns durchgeführte Transkriptionsassay zeigt, dass in Gegenwart beider Proteine zusätzliche Pausen während der Transkription vorhanden sind, welche möglicherweise einen Einfluss auf die Terminationseffizienz haben. Ein wichtiger Schritt hierbei könnte sein, dass NusA-AR2 die NusG-NTD von der RNAP ablöst und dadurch die Prozessivität der RNAP verringert wird, da Pausierungen begünstigt werden. Auch bei der Rho-abhängigen Termination konnte gezeigt werden, dass beide Proteine zusammen neue Terminationsstellen erzeugen, welche bei den einzelnen Faktoren nicht vorhanden sind (Muteeb *et al.*, 2012). Um die Bedeutung der Interaktion zwischen NusA-AR2 und NusG-NTD zu klären sind jedoch weitere Experimente nötig.

3.3 Zusammenbau, Reinigung und Aktivitätstest der RNAP und ihrer Untereinheiten

Die RNAP ist das zentrale Enzym der Transkription, da sie die Synthese der RNA katalysiert. Bakterielle RNAPs bestehen aus den fünf Untereinheiten $\alpha_2\beta\beta'\omega$ (Ebright, 2000). Bisher sind mehrere Kristallstrukturen unterschiedlicher Organismen bekannt (zusammengefasst in Sekine *et al.*, 2012). Diese geben allerdings keinen Aufschluss über Dynamiken, schwache Interaktionen und Domänenbewegungen, welche für die Funktion der RNAP jedoch essentiell sind. Mittels NMR-Spektroskopie könnten diese Fragestellungen zwar untersucht werden, aber aufgrund der Größe der RNAP (390 kDa) ist dies mit traditionellen NMR-Methoden nicht möglich. Allerdings können in solch großen Systemen Methylgruppen als NMR-Sonden verwendet werden. Hierbei liegt die gesamte RNAP deuteriert vor und nur die Methylgruppen der Isoleucine, Leucine und Valine sind ^{1}H, ^{13}C markiert (Methylgruppenmarkierung). Da diese Methylgruppen auch bei großen Molekülen ausreichend frei rotieren, können [^{1}H, ^{13}C]-Korrelationsspektren aufgenommen werden. Um auch selektiv eine bestimmte RNAP-Untereinheit methylgruppenmarkieren zu können wurden in Einzelarbeit C die Gene der einzelnen Untereinheiten getrennt voneinander kloniert und exprimiert und anschließend die RNAP funktionsfähig *in vitro* assembliert. Die übrigen Untereinheiten tragen somit keine Signale zum NMR-Spektrum bei (Abb. 15A). Zusätzlich hierzu wurden die einzelnen Untereinheiten einzeln kloniert, exprimiert und gereinigt. Für die Produktion und Reinigung der α- und β-Untereinheiten war vorwiegend ich verantwortlich, während meine Kollegin Johanna Drögemüller hauptsächlich an der β‘- und ω-Untereinheiten arbeitete.

Für die *in vitro*-Rekonstitution der *E. coli* RNAP wurden die Zellpellets der Bakterienkulturen mit den einzelnen überproduzierten Untereinheiten in denaturierendem Puffer vereinigt und aufgeschlossen. Der Zusammenbau der Untereinheiten erfolgte durch Entfernung des Harnstoffs mittels stufenweiser Dialyse. Anschließend wurde der Komplex mittels Affinitäts- und Größenausschlusschromatographie (SEC) gereinigt (Abb. 15B). Die assemblierte RNAP eluierte von der SEC-Säule in zwei

Elutionsmaxima (SEC-Peak 1 und 2). Diese entsprachen umgerechnet molekularen Masse von 980 kDa (SEC-Peak 1) und 507 kDa (SEC-Peak 2). Eine Analyse der Fraktionen mittels Natriumdodecylsulfat-Polyacrylamidgelelektrophorese (SDS-PAGE) zeigte, dass in beiden Peaks alle RNAP-Untereinheiten enthalten waren. Allerdings waren in der Probe des SEC-Peaks 1 viele Verunreinigungen vorhanden und die errechnete molekulare Masse des 2. SEC-Peaks von 507 kDa ließ vermuten, dass es sich hierbei um die korrekt assemblierte RNAP handelte. Bestätigt wurde dies durch eine Referenzprobe mit *in vivo* assemblierter RNAP ($RNAP^{nativ}$), welche auf der Höhe des zweiten SEC-Peaks eluierte. Hierbei handelt es sich um eine gereinigte RNAP, bei der sich alle Untereinheiten auf einem Plasmid befinden. Auch die CD-Spektren von $RNAP^{nativ}$ und einer Probe aus dem 2. SEC-Peak waren nahezu identisch, wohingegen das Spektrum eines Aliquots aus dem 1. SEC-Peak deutliche Unterschiede aufwies (siehe Einzelarbeit C, Abb. 1).

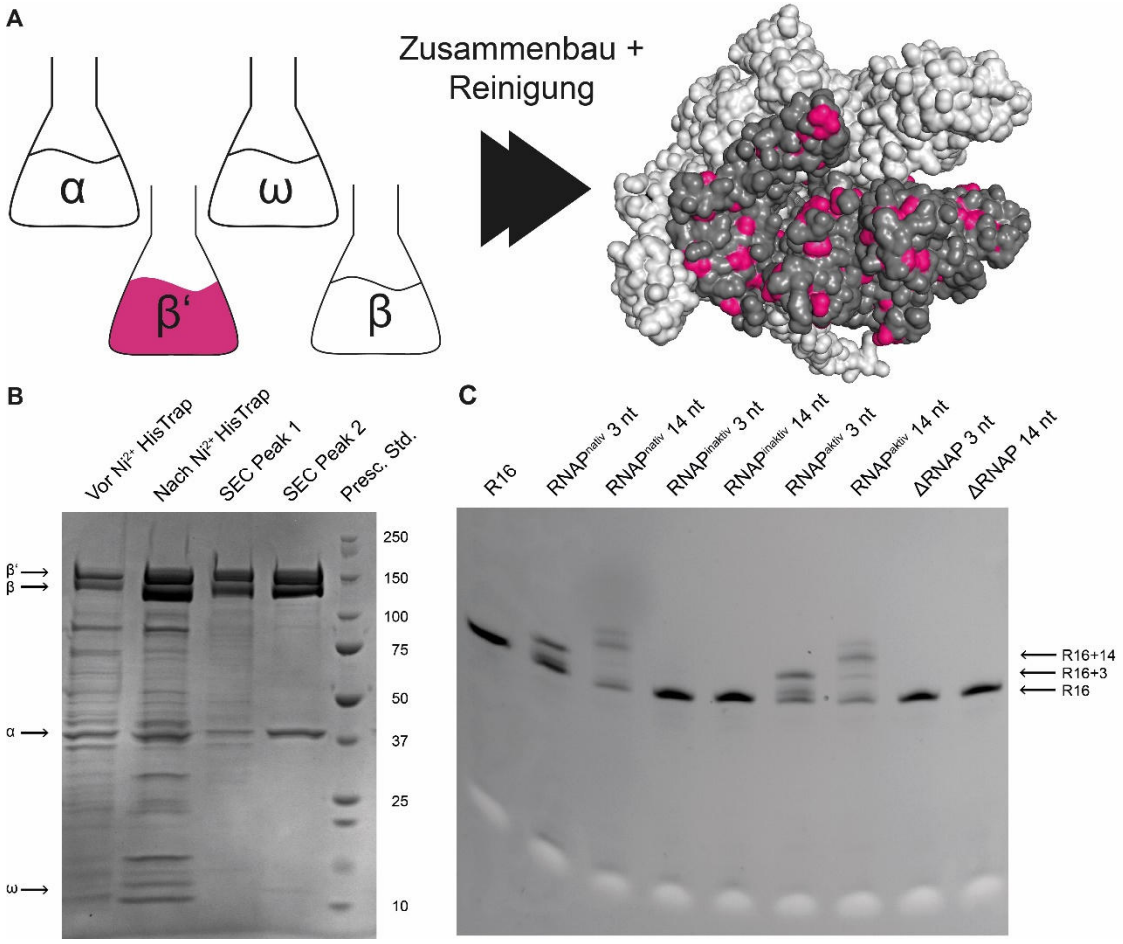


Abbildung 15: Reinigung und Aktivitätstest der *in vitro* assemblierten RNAP. A) Gezeigt ist ein Schema zur Herstellung aktiver RNAP mit methylgruppenmarkierter β'-Untereinheit. Die Oberfläche der RNAP (PDB-Code: 4KMU, Molodtsov *et al.*, 2013) ist in weiß gezeigt. Die β'-Untereinheit ist grau und die Ile, Leu und Val sind pink hervorgehoben. **B)** 4-20 %iges (w/v) SDS-Gradientenpolyacrylamidgel der RNAP-Reinigung nach Färbung mit Coomassie Blau. Vor Ni^{2+} HisTrap: lösliche Fraktion nach der Assemblierung; Nach Ni^{2+} HisTrap: Vereinte Fraktionen nach der Ni^{2+}-Affinitätschromatographie. SEC Peak 1: 1. Peak der SEC; SEC Peak 2: 2. Peak der SEC; Presc. Std.: Prescision Plus Protein Standard. Auf jede Bahn wurden 2 µg Protein aufgetragen. Die Molekulargewichtsgrößen des Standards, sowie die Laufhöhe der RNAP-Untereinheiten sind an den Seiten des Gels notiert. **C)** RNAP-Aktivitätstest, 20 %iges (w/v) SDS-Polyacrylamidgel. Es wurde entweder ATP und CTP (Bahn 2, 4, 6, 8) oder ATP, CTP und GTP (Bahn 3, 5, 7, 9) zugegeben, um eine Verlängerung der R16 um drei bzw. 14 nt zu ermöglichen. Die Negativkontrolle (ΔRNAP) wurde wie alle anderen Proben behandelt, allerdings wurde keine RNAP zugegeben. Auf Bahn 1 wurde unbehandelte R16 aufgetragen. Auf jede Bahn wurden 3 pmol RNA aufgetragen. Die Pfeile zeigen die Laufhöhe der R16 sowie der elongierten RNAs an.

Um den korrekten Zusammenbau der RNAP zu überprüfen wurde ein Aktivitätstest durchgeführt. Hierzu wurde mittels eines DNA-Matrizenstranges, eines nicht kodierenden DNA-Stranges und eines kurzen, fluoreszenzmarkierten RNA-Stranges (R16) der TEC ausgebildet. Anschließend wurde die Transkription durch Zugabe von ATP und CTP oder durch Zugabe von ATP, CTP und GTP gestartet, wodurch die RNA um drei bzw. 14 Nukleotide in Gegenwart von aktiver RNAP verlängert werden konnte. Die Auswertung des Aktivitätstests erfolgte durch SDS-PAGE und durch Detektion der RNA mittels Fluoreszenz. Die *in vitro* assemblierte RNAP des zweiten SEC-Peaks war ähnlich aktiv wie $RNAP^{nativ}$, während es bei der Zugabe von RNAP aus dem ersten SEC-Peak zu keiner Verlängerung des RNA-Strangs kam (Abb. 15C). Der aktive Teil der *in vitro* assemblierten RNAP wird daher im Folgenden als $RNAP^{aktiv}$, der inaktive Teil als $RNAP^{inaktiv}$ bezeichnet.

Im Vergleich zu bisher publizierten Protokollen für die *in vitro*-Assemblierung der RNAP, hat die hier etablierte Methode den Vorteil, dass die Untereinheiten vorher nicht separat gereinigt werden müssen und daher Material und Arbeitszeit gespart wird (Heil und Zillig, 1970; Borukhov und Goldfarb, 1993; Tang *et al.*, 1995; Palm *et al.*, 1975). Die Ausbeute und Reinheit ist vergleichbar mit den bereits publizierten Protokollen. Allerdings wurden die vorherigen Assemblierungen ohne die ω-Untereinheit durchgeführt. Diese wird durch das Gen *rpoZ* kodiert und ist weder für die Zellviabilität, noch für die Funktion der RNAP essentiell (Gentry *et al.*, 1991). Allerdings kommt es zu einer Erhöhung der RNAP-Aktivität, wenn die Assemblierung in Gegenwart der ω-Untereinheit durchgeführt wird, da sie die Faltung der β'-Untereinheit und die Assemblierung von $\alpha_2\beta$ mit β'ω fördert (Ghosh *et al.*, 2001; Mukherjee *et al.*, 1999). Der Zusammenbau in Gegenwart der ω-Untereinheit hat daher vermutlich weniger falsch gefaltete oder falsch assemblierte RNAP zur Folge. Ein weiterer Vorteil der hier gezeigten Reinigungsstrategie ist, dass der Initiationsfaktor σ nicht benötigt wird. Somit entspricht die gereinigte RNAP der Zusammensetzung während der Transkriptionselongation und zusätzliche Schritte zur Abtrennung des σ-Faktors sind nicht notwendig. Durch die Separation von inaktiven und fehlgefalteten Varianten durch die SEC, wird außerdem sichergestellt, dass nur aktive RNAP vorhanden ist. Dieser Reinigungsschritt wurde in den meisten vorherigen Protokollen nicht durchgeführt.

3.4 Untersuchung der RNAP mittels NMR-Spektroskopie

Zunächst wurde ein [^{1}H, ^{13}C]-TROSY-*Heteronuclear Multiple Quantum-Coherence* (HMQC) Spektrum der deuterierten $RNAP^{nativ}$ aufgenommen, wobei die Methylgruppen der Aminosäurereste Isoleucin, Leucin und Valin aller Untereinheiten ^{1}H, ^{13}C markiert waren (Abb. 16). Das TROSY-HMQC ist eine Abwandlung des HSQC-Experiments. Diese Pulsfolge liefert für große, methylgruppenmarkierte Proteine bessere Spektren (Tugarinov *et al.*, 2003).

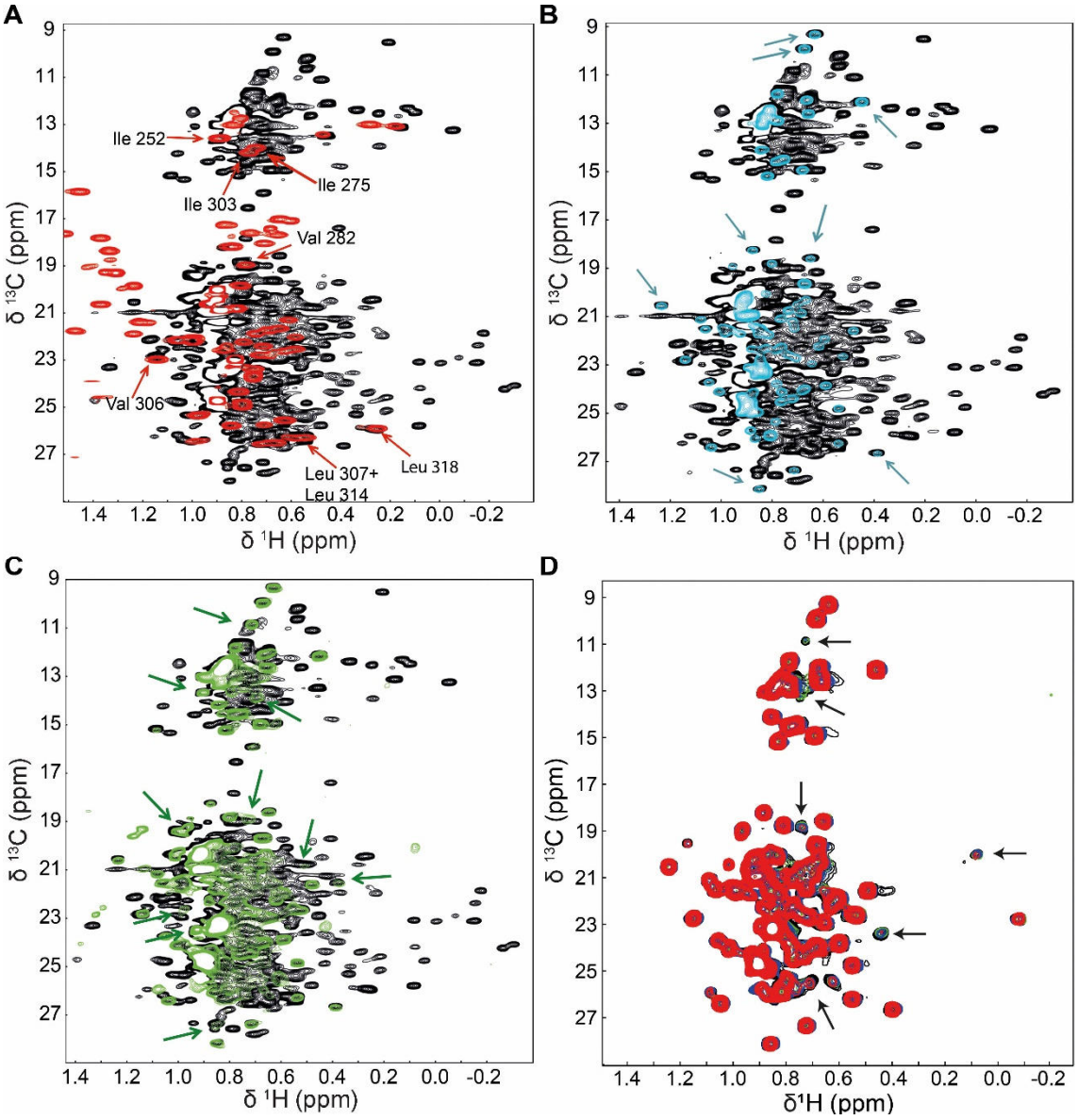


Abbildung 16: NMR-Untersuchungen an der RNAP. Gezeigt sind C-H Korrelationsspektren von Methylgruppen der RNAP^{nativ}, der RNAP αCTD, von der freien β'-Untereinheit und der β'-Untereinheit eingebaut in die RNAP. Bei RNAP^{nativ} und β' sind diese methylgruppenmarkiert, wohingegen bei β' in RNAP nur β' methylgruppenmarkiert ist, während die restliche RNAP deuteriert ist. Die RNAP αCTD war komplett ^{15}N, ^{13}C markiert. **A)** [^{1}H, ^{13}C]-HMQC-Spektren von RNAP^{nativ} (schwarz, 30 µM) und [^{1}H, ^{13}C]-HSQC-Spektrum von RNAP αCTD (rot, 700 µM). Einige zugeordnete Signale sind durch rote Pfeile markiert. **B)** [^{1}H, ^{13}C]-HMQC-Spektren von RNAP^{nativ} (schwarz) und β' (cyan, 2 µM). Die Peaks mit denselben chemischen Verschiebungen in RNAP^{nativ} und freiem β' sind durch blaue Pfeile markiert. **C)** [^{1}H, ^{13}C]-HMQC-Spektren von RNAP^{nativ} (schwarz) und β' in RNAP (grün). Die grünen Pfeile zeigen Signale an, welche im Spektrum von isolierter β' nicht vorhanden sind. **D)** [^{1}H, ^{13}C]-HMQC-Spektren von β' vor (schwarz) und nach Zugabe von unmarkiertem NusG-NTD in einem Verhältnis von 1:1 (grün), 1:2 (blau) und 1:10 (rot). Die Pfeile zeigen verschwindende Signale an.

Das Spektrum der RNAP^{nativ} zeigt eine breite Dispersion der Methylgruppensignale mit chemischen Verschiebungen der Protonenresonanzfrequenzen im Bereich von 1,4 bis -0,3 ppm. Dies deutet auf ein gefaltetes Protein hin. Es ist ersichtlich, dass es aufgrund der Vielzahl an Signalen (230 Ile, 349 Leu, 287 Val) zu Überlagerungen im Spektrum kommt, vor allem im Protonenresonanzfrequenzbereich von 1,0 bis 0,5 ppm. Durch die Überlagerung dieses Spektrums mit dem Metyhlgruppenbereich des [^{1}H, ^{13}C]-HSQC-Spektrums der RNAP αCTD konnten einige Signale direkt der RNAP αCTD

zugeordnet werden (Abb. 16A). Dies war möglich, da viele Signale identische chemische Verschiebungen aufweisen und die Signale der RNAP αCTD im RNAP^{nativ} Spektrum sehr intensiv sind, da die RNAP αCTD flexibel an die restliche RNAP gebunden ist. Außerdem war die Zuordnung der isolierten RNAP αCTD bereits bekannt (Schweimer *et al.*, 2011). Auf analoge Weise wurden die zur β'-Untereinheit gehörenden Signale zugeordnet. Hierzu wurde β' methylgruppenmarkiert gereinigt, ein [^{1}H, ^{13}C]-TROSY-HMQC-Spektrum aufgenommen und mit dem Spektrum der RNAP^{nativ} überlagert (Abb. 16B). Die Signale der β'-Untereinheit sind ebenfalls dispergiert und einige Signale der RNAP^{nativ} können eindeutig β' zugewiesen werden, da die Resonanzfrequenzen in beiden Spektren nahezu identisch sind. Da für die β'-Untereinheit keine Zuordnung vorhanden ist, konnte noch nicht bestimmt werden, um welche Sequenzpositionen es sich hierbei handelt. Das [^{1}H, ^{13}C]-TROSY-HMQC-Spektrum der methylgruppenmarkierten β'-Untereinheit in *in vitro* assemblierter RNAP liefert erheblich mehr Signale als das Spektrum der isolierten β'-Untereinheit. Hierbei sind alle Signale der isolierten β'-Untereinheit noch bei nahezu identischen Resonanzfrequenzen zu finden, aber die zusätzlichen Signale überlagern mit denen im RNAP^{nativ}-Spektrum (Abb. 16C). Mit Hilfe des Spektrums lassen sich daher weitere Signale der β'-Untereinheit zuordnen. Vermutlich kommen diese zusätzlichen Signale daher, dass sich die β'-Untereinheit im physiologischen Kontext befindet.

Um zu überprüfen, ob die isolierte, methylgruppenmarkierte β'-Untereinheit immer noch NusG-NTD binden kann, wurde unmarkierte, protonierte NusG-NTD zutitriert. Während die meisten Signalintensitäten unverändert blieben, kam es bei einigen Signalen zu signifikanten Intensitätsverlusten (Abb. 16D). Der Signalverlust beruht hierbei darauf, dass die angeregten Methylgruppen der β'-Untereinheit Magnetisierung auf die benachbarten Protonen der NusG-NTD übertragen können wenn NusG-NTD gebunden ist. Dadurch wird die transversale Relaxation beschleunigt und Signalintensität geht verloren. In der Bindungsstelle der NusG-NTD an die β'CH sind zwei Ile und zwei Leu positioniert (Martinez-Rucobo *et al.*, 2011). Diese sollten zwei Signale in der Ile-Region ($\delta(^{13}C)$ = 9-16 ppm) und vier Signale in der Val/Leu-Region ($\delta(^{13}C)$ = 17-29 ppm) zum Spektrum beitragen. Da bei der Zugabe der NusG-NTD diese Anzahl an Signalen einen signifikanten Intensitätsverlust erleidet, kann davon ausgegangen werden, dass die NusG-NTD an die richtige Stelle der β'-Untereinheit bindet und die β'-Untereinheit die korrekte Faltung angenommen hat.

Der Zusammenbau der RNAP aus den einzelnen Untereinheiten und die Aufnahme der Methylgruppenspektren ist ein erster wichtiger Schritt zur Untersuchung der RNAP mittels NMR-Spektroskopie. Basierend auf diesen Messungen können weitere RNAP-Konstrukte geplant werden, um Bereiche der einzelnen Untereinheiten zuzuordnen. Außerdem können Bindungen gemessen und strukturelle Änderungen der RNAP bestimmt werden. Besonders interessant sind diese Messungen, da erstmals gezeigt wurde, dass bei methylgruppenmarkierten Proteinkomplexen das zu messende Protein nicht mehrfach in dem zu untersuchenden Komplex vorkommen muss, um qualitativ gute Spektren zu erhalten.

3.5 Bestimmung der mit einem Transkriptionsfaktor interagierenden RNAP-Untereinheit

Viele Prozesse und Regulationen der Transkription basieren auf direkten Interaktionen regulatorischer Faktoren mit der RNAP. Beispiele hierfür sind Antibiotika wie Rifamycine (Hartmann *et al.*, 1967) oder Myxopyronine (Irschik *et al.*, 1983), aber auch die Transkriptionsfaktoren NusG, NusA und NusE (Mooney *et al.*, 2009a; Mason und Greenblatt, 1991; Mah *et al.*, 1999). Um die molekulare Basis der resultierenden Effekte zu verstehen, ist es nötig die Bindungsstelle an die RNAP zu ermitteln. Während bekannt war, dass NusG-NTD an die β'CH und den βGL bindet (Martinez-Rucobo *et al.*, 2011; Sevostyanova *et al.*, 2011) und NusA-AR2 mit der RNAP αCTD wechselwirkt (Mah *et al.*, 1999) wurde bei NusA-NTD lediglich vermutet, dass die Bindung in der Nähe des RNA-Austrittskanals an die βFTH erfolgt (Yang *et al.*, 2009; Ha *et al.*, 2010; Yang und Lewis, 2010). Für NusE war keine Bindestelle an die RNAP bekannt. In Einzelarbeit C wurde eine Methode entwickelt, um zu bestimmen, welche Untereinheit der RNAP mit dem jeweiligen Nus-Faktor interagiert. Zur Validierung der Methode sollten die bekannten bzw. vorgeschlagenen Bindungsstellen von NusG-NTD, NusA-AR2 und NusA-NTD bestätigt und die mit NusE interagierende Untereinheit bestimmt werden. Gleichzeitig konnte hierdurch überprüft werden, ob die isoliert gereinigten RNAP-Untereinheiten korrekt gefaltet sind.

Als erstes wurde die Sekundärstruktur der einzelnen RNAP-Untereinheiten mittels CD-Spektroskopie analysiert (Einzelarbeit C Abb. 1). Für alle Untereinheiten mit Ausnahme von ω wurden Spektren erhalten, die für gefaltete Proteine typisch sind. Von ω war jedoch bereits bekannt, dass es alleine keine definierte Sekundärstruktur annimmt und lediglich der N-Terminus eine α-helikale Struktur ausbildet (Ghosh *et al.*, 2001; Greenfield und Fasman, 1969). Um die interagierende RNAP-Untereinheit zu bestimmen, wurden die zu untersuchenden Transkriptionsfaktoren bzw. einzelne Domänen ^{15}N isotopenmarkiert gereinigt, mit den Untereinheiten der RNAP titriert und durch 1D oder 2D [^{1}H, ^{15}N]-HSQC-Spektren die Signalintensitäten bzw. die Änderungen der chemischen Verschiebungen beobachtet. Für NusG-NTD, NusA-AR2 und NusA-NTD konnten hierdurch die prognostizierten Bindungsstellen bestätigt werden, wodurch die Methode validiert wurde. Die Ergebnisse für NusE sind in Abb. 17 gezeigt. Für diese Messungen wurde aus Stabilitätsgründen der Komplex NusE$^{\Delta}$:NusB verwendet. In NusE$^{\Delta}$ ist aus Löslichkeitsgründen die Ribosomenbindungs-schleife durch ein Serin ersetzt. Bei Zugabe von RNAP^{nativ} zu ^{15}N-NusE$^{\Delta}$:NusB kam es zu einem vollständigen Signalverlust, was auf eine Bindung der RNAP an NusE$^{\Delta}$:NusB hinweist. Der Signalverlust konnte auch bei Zugabe der isolierten β-Untereinheit beobachtet werden. Es kommt also zu einer Komplexbildung zwischen der β-Untereinheit und NusE$^{\Delta}$:NusB. In Anwesenheit von β' ist die Signalabnahme deutlich geringer. Dies ist vermutlich auf eine unspezifische Wechselwirkung zurückzuführen. Bei Zugabe der α- bzw. ω-Untereinheit sind kaum Veränderungen in den Spektren erkennbar. Diese sind also nicht an der Bindung beteiligt. NusE$^{\Delta}$:NusB bindet daher vermutlich an die β-Untereinheit.

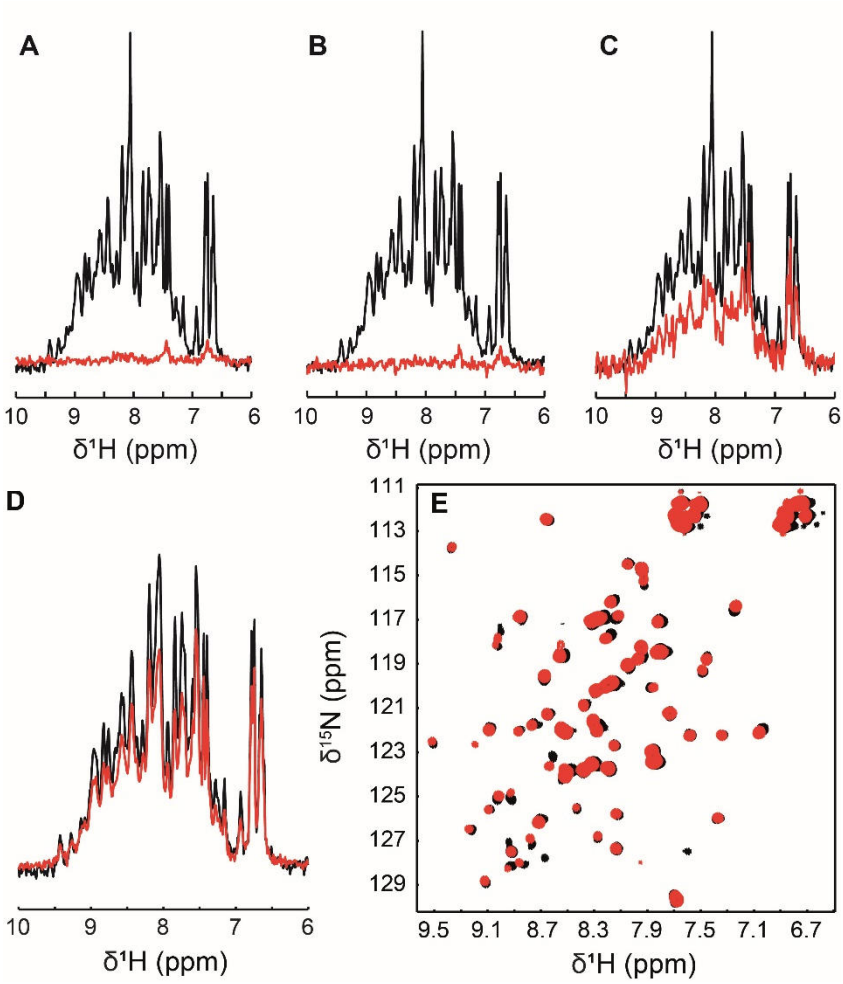


Abbildung 17: Interaktion von NusE$^{\Delta}$:NusB mit der RNAP und ihrer Untereinheiten. 1D [^{1}H, ^{15}N]-HSQC-Spektren der Amidregion von ^{15}N-NusE$^{\Delta}$:NusB in Abwesenheit (schwarz) und in Gegenwart (rot) von äquimolaren Konzentrationen der **A)** RNAP^{nativ}, **B)** β-Untereinheit, **C)** β'-Untereinheit, **D)** α-Untereinheit. **E)** 2D [^{1}H, ^{15}N]-HSQC-Spektren von ^{15}N-NusE$^{\Delta}$:NusB in Abwesenheit (schwarz) und in Gegenwart (rot) von äquimolaren Konzentration der ω-Untereinheit.

Durch diese Methode wurde bestätigt, dass NusG-NTD sowohl an die β- als auch an die β'-Untereinheit der RNAP bindet. Dies ist in Übereinstimmung damit, dass NusG-NTD die beiden Klammern der RNAP verknüpft und, aufgrund der stärkeren DNA-Bindung, die Prozessivität der RNAP erhöht (Sevostyanova *et al.*, 2011). Die Wechselwirkung der NusA-AR2 mit der RNAP αCTD konnte ebenfalls verifiziert werden. Die Ergebnisse für NusA-NTD bekräftigten die Prognose, dass die Bindung an die β-Untereinheit erfolgt. Gleichzeitig wurde hierdurch bestätigt, dass die isolierten RNAP-Untereinheiten in dem Bereich der Bindungsstellen die korrekte Faltung annehmen. Für NusE$^{\Delta}$ konnte gezeigt werden, dass es mit der β-Untereinheit der RNAP interagiert. Obwohl die genaue Bindungsstelle weiterhin unbekannt ist, können zukünftige Experimente zur Untersuchung der RNAP:NusE-Interaktion zielgerichteter und effektiver geplant und durchgeführt werden. Theoretisch könnte durch die Bindung von NusE an die RNAP die Transkription direkt mit der Translation verknüpft sein, da NusE als S10 ein Teil des Ribosoms ist. Dies ist jedoch unwahrscheinlich, da der Komplex aus RNAP und Ribosom ziemlich starr und die Genexpression vermutlich beeinträchtigt

wäre. Wahrscheinlich spielt die Interaktion eine Rolle bei der Transkriptionsantitermination, wie bereits schon früher vermutet wurde (Mason und Greenblatt, 1991).

Generell ist diese einfache, schnelle Messmethode, die auf konventionellen NMR-Experimenten beruht, geeignet um festzustellen mit welcher Domäne der RNAP kleinere Moleküle, wie beispielsweise Antibiotika und kleinere Proteine, interagieren.

3.6 Bestimmung der RNAP-Bindungsflächen von NusG-NTD, NusA-NTD und NusE

Die Nus-Faktoren haben einen großen Einfluss auf die Transkription und spielen besonders bei der Elongation und der Termination eine wichtige Rolle (Mooney *et al.*, 1998). Dennoch waren die RNAP-Bindungsflächen der Nus-Faktoren größtenteils unbekannt. Aufgrund einer Kristallstruktur war die Bindungsfläche von NusG-NTD an die β'CH aufgeklärt und die Bindungsfläche an den βGL wurde auf Basis eines Homologiemodells prognostiziert (Martinez-Rucobo *et al.*, 2011; Sevostyanova *et al.*, 2011). Für NusA-NTD konnte durch Cryoelektronenmikroskopie gezeigt werden, an welche genaue Stelle der RNAP es bindet. Allerdings war die Auflösung mit 22 Å relativ niedrig und die Bindungsfläche auf Seite der NusA-NTD konnte nicht eindeutig bestimmt werden (Yang *et al.*, 2009). Durch Peptidschnitte mittels FeBABE (Fe(III) Komplex von (S)-2-[4-(2-Bromoacetamido)benzyl]-ethylendiamintetraacetat) und die zwei Varianten NusA(S29C) und NusA(S53C) konnte gezeigt werden, dass S29 in der Interaktionsfläche liegt, während S53 nicht an der Bindung beteiligt ist (Ha *et al.*, 2010). Hierdurch lässt sich die Orientierung der NusA-NTD allerdings nur grob abschätzen. Für NusE waren keinerlei Informationen über die Bindungsfläche bekannt. In Einzelarbeit D wurde daher eine Methode entwickelt, um die RNAP-Bindungsstellen der Nus-Faktoren detailliert zu bestimmen.

Für die Bestimmung der Bindungsflächen von den Nus-Faktoren an die RNAP, wurden die Nus-Faktoren methylgruppenmarkiert und [^{1}H, ^{13}C]-Methylgruppen-TROSY-Spektren in D_2O aufgenommen. Anschließend wurde protonierte RNAP (in D_2O) zutitriert. Die Signale der an der RNAP-Bindung beteiligten Reste nehmen dabei stärker ab, da die Magnetisierung der Methylgruppen auf die Protonen der RNAP übertragen werden kann und betroffene Reste somit schneller relaxieren. Durch die Bestätigung der bereits bekannten Bindungsflächen von NusG-NTD wurde die Methode validiert. In Abb. 18A sind die Spektren der Titration von methylgruppenmarkierter NusA-NTD mit RNAP gezeigt. Zur Ermittlung der Bindungsfläche wurden die Reste in stark und schwach betroffen eingeteilt und auf der jeweiligen Struktur markiert (Abb. 18B). Da die Messungen lediglich Auskünfte über die Ile, Leu und Val geben, wurden zwei benachbarte Aminosäuren zusätzlich markiert, solange es sich nicht um ein nicht betroffenes Iso, Leu oder Val handelte.

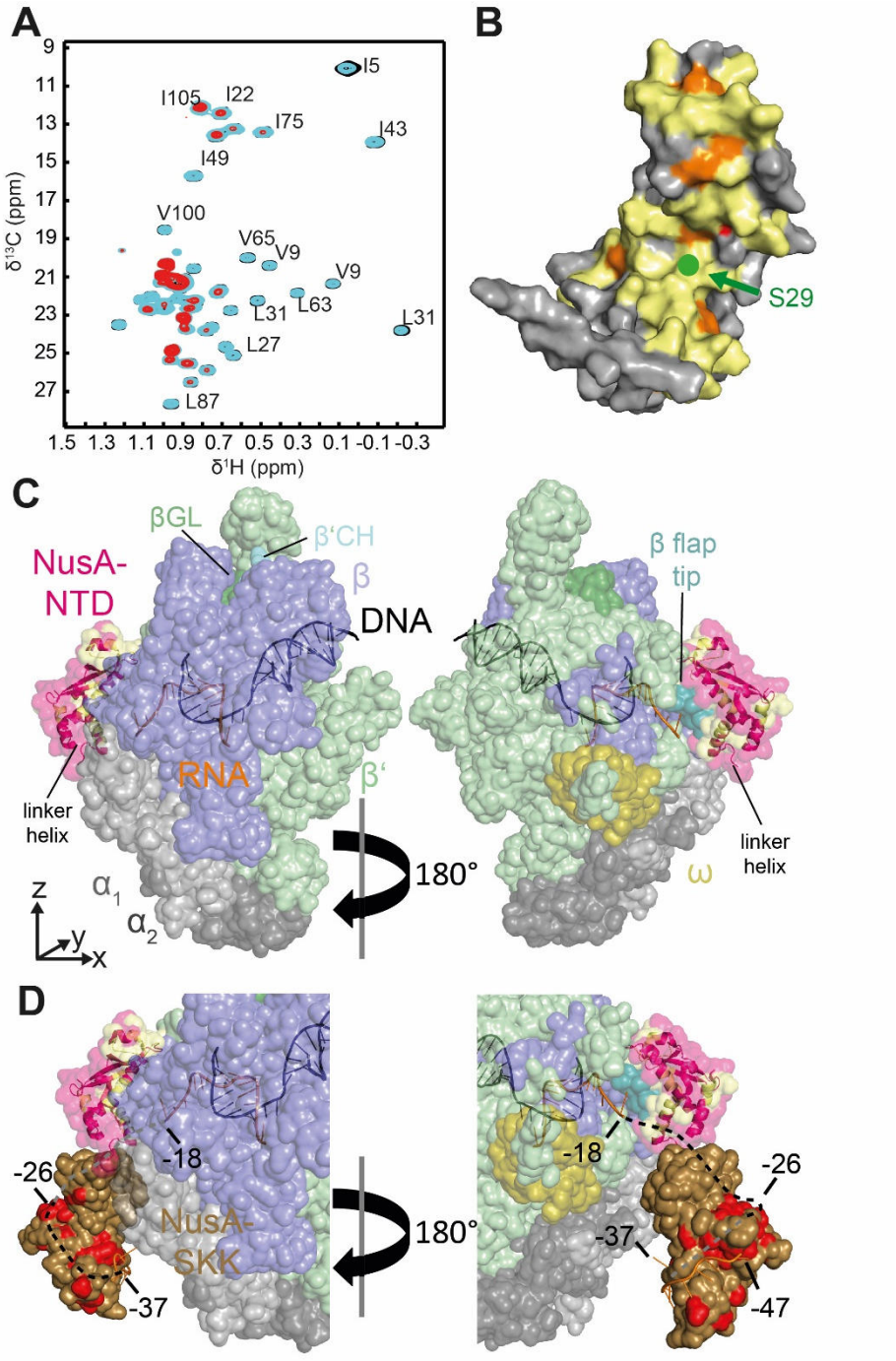


Abbildung 18: Modell für die Bindung der NusA-NTD an die RNAP. A) Titration der methylgruppenmarkierten NusA-NTD mit RNAP. Die Methyl-TROSY-Spektren von NusA-NTD in Abwesenheit (schwarz) und in Gegenwart der RNAP (1:1 molares Verhältnis, cyan; 1:2 molares Verhältnis, rot) sind gezeigt. Einige zugeordnete Signale sind beschriftet. **B)** Die bei der Titration betroffenen Reste sind auf der Struktur der NusA-NTD gekennzeichnet (PDB-Code: 2KWP, Einzelarbeit C). Stark betroffene Reste sind rot, leicht betroffene Reste orange eingefärbt. Wenn ein Rest betroffen war, wurden die zwei benachbarten Reste gelb eingefärbt, es sei denn es handelste sich um ein nicht betroffenes Ile, Leu oder Val. **C)** Modell des NusA-NTD:RNAP-Komplexes. NusA-NTD (pink) ist in Cartoon- und Oberflächendarstellung gezeigt. Die ermittelte Bindungsfläche ist hellgelb markiert. Das Modell ist die beste Lösung einer HADDOCK-Simulation von NusA-NTD mit elongierender *T. thermophilus* RNAP (Oberflächendarstellung; β, blau; β', hellgrün; β'CH, hellblau; βGL, dunkelgrün; α_1 und α_2, grau; RNA, orange; βFTH türkis; ω, gelb PDB-Code: 2O5I, Vassylyev *et al.*, 2007a). **D)** Bindung der naszierenden RNA durch NusA. Die Orientierung der NusA-NTD ist wie in C). Um die NusA-SKK mit gebundener RNA zu positionieren wurde *Tm*NusA (PDB-Code: 1L2F, Shin *et al.*, 2003) und *Mt*NusA (PDB-Code: 2ASB, Beuth *et al.*, 2005) mit der Struktur von *Ec*NusA-NTD überlagert. NusA-SKK ist in Oberflächendarstellung (braun) gezeigt und die durch die RNA-Bindung betroffenen Reste nach Schweimer *et al.* 2011 rot gekennzeichnet. RNA aus der Kristallstruktur von *Mt*NusA-SKK ist orange eingezeichnet. Die gestrichelte Linie zeigt einen möglichen Weg für die naszierende RNA. Die abgeschätzten Basennummern sind eingezeichnet.

Durch diese Messmethode wurden bei der NusG-NTD zwei betroffene Regionen an der Proteinoberfläche ermittelt (Abb. 19). Diese entsprechen den prognostizierten Bindungsflächen an die β‘CH und den βGL (Martinez-Rucobo *et al.*, 2011; Sevostyanova *et al.*, 2011). Hierdurch wird das Modell bestätigt, wonach NusG gleichzeitig an die β‘CH und den βGL binden kann und die Prozessivität der RNAP erhöht, indem die beiden Klammerhälften miteinander verknüpft werden und die Bindung an die DNA verstärkt wird (Sevostyanova *et al.*, 2011). Für NusA-NTD konnte eine zusammenhängende Bindungsfläche ermittelt werden (Abb. 18, 19). Die Ergebnisse der FeBABE-Messungen wurden durch unsere Messungen bestätigt (Ha *et al.*, 2010), da sich der Rest S29 innerhalb der Bindungsfläche befindet, während der Rest S53 außerhalb liegt. Da durch unsere Methode die Bindungsfläche genauer bestimmt wurde, ergab sich, dass der negativ geladene Kopf der NusA-NTD ebenfalls an der RNAP-Bindung beteiligt ist und NusA-NTD im Vergleich zu dem vorgeschlagenen Modell von Ha *et al.,* 2010 um ca. 100° um die z-Achse gedreht ist. Basierend auf den betroffenen Resten wurde mittels HADDOCK (de Vries *et al.*, 2010) ein Bindungsmodell der NusA-NTD an die RNAP erstellt (Abb. 18C). Zusammen mit der bereits bekannten RNA-Bindungsflächen an die NusA-SKK (Schweimer *et al.*, 2011), konnte ein Model erstellt werden, wie die naszierende RNA durch die NusA-NTD zur NusA-SKK gefädelt wird und sich um diese herumwindet (Abb. 18D).

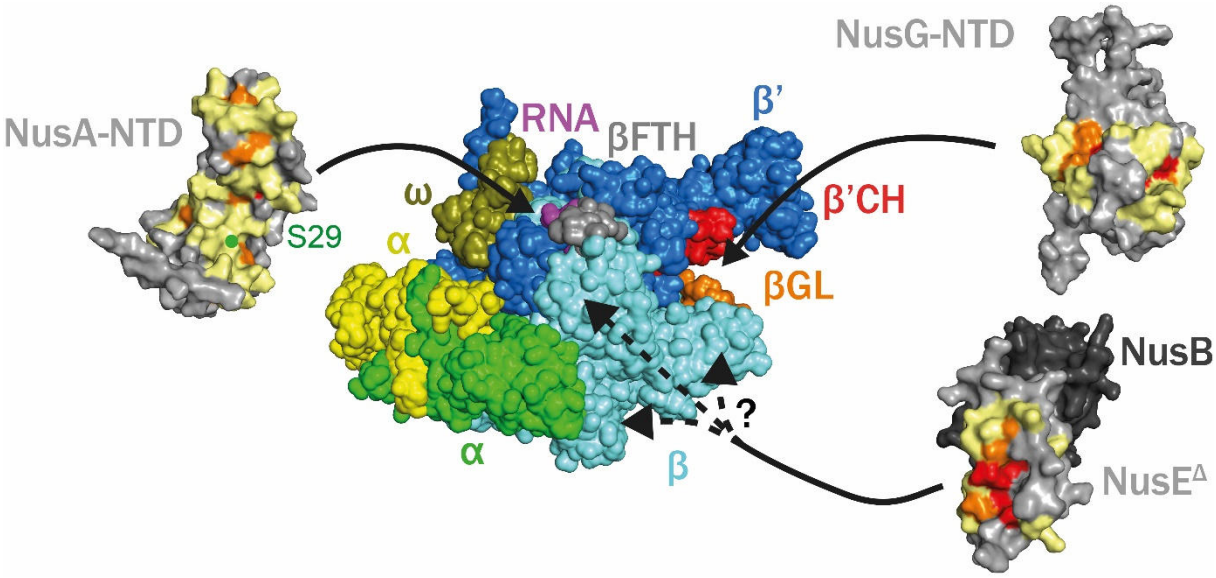


Abbildung 19: Bindungsstellen der Nus-Faktoren an die RNAP. Die Struktur der RNAP ist zentral in Oberflächendarstellung gezeigt (PDB-Code: 2O5I). α_1: gelb α_2: grün; β: hellblau, β‘: dunkelblau; ω: ocker; β’CH: rot; βGL: orange; βFTH: grau; RNA: lila. Die Strukturen der NusA-NTD (PDB-Code: 2KWP, Einzelarbeit C), NusG-NTD (PDB-Code: 2K06, Mooney *et al.*, 2009b) und des $NusE^{\Delta}$:NusB-Heterodimers (PDB-Code: 3D3B, Luo *et al.*, 2008) sind ebenfalls in Oberflächendarstellung gezeigt. Die bei der Titration mit der RNAP stark betroffenen Reste sind in rot gekennzeichnet, die mäßig Betroffenen in orange. In gelb sind benachbarte Reste von betroffenen Resten markiert, sofern sie nicht selbst ein nicht betroffenes Ile, Val oder Leu waren. Die Pfeile zeigen an welche Stelle der RNAP der jeweilige Faktor bindet. Während für NusA-NTD und NusG-NTD die Bindungsflächen auf Seiten der RNAP gut bestimmt sind, kann für $NusE^{\Delta}$ nur gesagt werden, dass es an die β-Untereinheit bindet.

Für $NusE^{\Delta}$ konnte ebenfalls eine einheitliche Bindungsfläche an die RNAP bestimmt werden (Abb. 19). Im $NusE^{\Delta}$:NusB-Heterodimer ist diese Fläche frei zugänglich, jedoch überlagert sie mit der Bindungsfläche der NusG-CTD (Burmann *et al.*, 2010). Durch HSQC-Verdrängungsexperimente wurde gezeigt, dass eine gleichzeitige Bindung der NusG-CTD und RNAP nicht möglich ist und die Affinitäten von NusG-CTD und RNAP an NusE ähnlich sind (Einzelarbeit D, Abb. 3). Daher liegt für die Bindung von NusE an NusG-CTD oder RNAP ein Gleichgewicht vor, zumindest wenn keine weiteren Einflüsse oder Faktoren vorhanden sind. Mögliche Faktoren könnten beispielsweise weitere Nus-Faktoren oder bestimmte Sequenzen der naszierenden RNA sein. Die Verdrängungsexperimente zeigen außerdem, dass die ermittelte Bindungsfläche an die RNAP korrekt ist.

Generell kann durch die vorgestellte Messmethode die RNAP-Bindungsfläche von jedem beliebigen Protein bestimmt werden. Hierbei sind keine aufwendigen Klonierungen oder Mutationen nötig und die Bindung kann am Wildtypprotein gemessen werden. Prinzipiell ist diese Methode auch für andere Systeme geeignet bei dem ein kleiner Bindungspartner mit einem supramolekularen Komplex interagiert. Im Bereich der Transkription könnten die Bindungsstellen der Nus-Faktoren an den Rho-Faktor analog bestimmt werden.

4. Abkürzungsverzeichnis

A. aeolicus	*Aquifex aeolicus*
*Aa*NusG	NusG aus *Aquifex aeolicus*
AR	*Acidic repeat*
β'CH	β' *clamp helices* (β' Klammerhelices)
βFTH	β *flap tip helix* (β Klappenspitzehelix)
βGL	β *gate loop* (β Pfortenschleife)
bp	Basenpaare
ATP	Adenosintriphosphat
CAP	*catabolite gene activator protein*
CD	Circulardichroismus
ChIP-chip-Experimente	Chromatin-Immunopreziptiations-Experimente
CTP	Cytidintriphosphat
DNA	Desoxyribonukleinsäure
*Ec*NusG	NusG aus *Escherichia coli*
E. coli	*Escherichia coli*
GTP	Guanosintriphosphat
HMQC	*Heteronuclear Multiple Quantum-Coherence*
HSQC	*Heteronuclear single quantum coherence*
kDa	Kilodalton
KH	K homologe Domäne
*Mt*NusG	NusG aus *Mycobacterium tuberculosis*
M. tuberculosis	*Mycobacterium tuberculosis*
NMR	*Nuclear magnetic resonance spectroscopy*
NOE	Nuklearer Overhauser Effekt
nt	Nukleotide
NTD	Aminoterminale Domäne
NTP	Ribonukleosidtriphosphat
Nus A, B, E, G	*N utilization substance A, B, E, G*
$NusE^{\Delta}$	NusE bei dem die Reste 46-67 durch ein Serin ersetzt wurden
nut	*N utilization site*
ops-DNA	*operon polarity suppressor* DNA
PP_i	Pyrophosphat
RNA	Ribonukleinsäure
RNAP	RNA-Polymerase
RNAP αCTD	Carboxy-terminale Domäne der α-Untereinheit der RNAP

RNAP αNTD	Amino-terminale Domäne der α-Untereinheit der RNAP
$RNAP^{aktiv}$	Aktiver Anteil der *in vitro* assemblierten RNAP
$RNAP^{inaktiv}$	Inaktiver Anteil der *in vitro* assemblierten RNAP
$RNAP^{nativ}$	Aus einem Plasmid gereinigte RNAP
rrn	Ribosomale RNA
rut	*Rho utilization site*
S1	S homologe Domäne 1
SDS-PAGE	Natriumdodecylsulfat-Polyacrylamidgelelektrophorese
SEC	Größenausschlusschromatographie
TEC	Transkriptionselongationskomplex
T. maritima	*Thermatoga maritima*
*Tm*NusG	NusG aus *Thermatoga maritima*
T. thermophilus	*Thermus thermophilus*
*Tt*NusG	NusG aus *Thermus thermophilus*
TROSY	*transverse relaxation optimized spectroscopy*

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

Amero C, Asuncion Dura M, Noirclerc-Savoye M, Perollier A, Gallet B, Plevin MJ, Vernet T, Franzetti B und Boisbouvier J (2011) A systematic mutagenesis-driven strategy for site-resolved NMR studies of supramolecular assemblies. *J Biomol NMR* **50:** 229-236

Arndt KM, Chamberlin MJ (1990) RNA chain elongation by *Escherichia coli* RNA polymerase. Factors affecting the stability of elongating ternary complexes. *J Mol Biol* **213:** 79-108

Bailey MJA, Hughes C und Koronakis V (1997) RfaH and the ops element, components of a novel system controlling bacterial transcription elongation. *Mol Microbiol* **26:** 845-851

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

Bar-Nahum G, Nudler E (2001) Isolation and characterization of sigma(70)-retaining transcription elongation complexes from *Escherichia coli*. *Cell* **106:** 443-451

Barne KA, Bown JA, Busby SJ und Minchin SD (1997) Region 2.5 of the *Escherichia coli* RNA polymerase sigma70 subunit is responsible for the recognition of the 'extended-10' motif at promoters. *EMBO J* **16:** 4034-4040

Basu RS, Warner BA, Molodtsov V, Pupov D, Esyunina D, Fernandez-Tornero C, Kulbachinskiy A und Murakami KS (2014) Structural Basis of Transcription Initiation by Bacterial RNA Polymerase holoenzyme. *J Biol Chem* **289:** 24549-59

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

Beuth B, Pennell S, Arnvig KB, Martin SR und Taylor IA (2005) Structure of a *Mycobacterium tuberculosis* NusA-RNA complex. *EMBO J* **24:** 3576-3587

Bonin I, Muhlberger R, Bourenkov GP, Huber R, Bacher A, Richter G und Wahl MC (2004) Structural basis for the interaction of *Escherichia coli* NusA with protein N of phage lambda. *Proc Natl Acad Sci* **101:** 13762-7

Borukhov S, Goldfarb A (1993) Recombinant *Escherichia coli* RNA polymerase: purification of individually overexpressed subunits and *in vitro* assembly. *Protein Expr Purif* **4:** 503-511

Borukhov S, Lee J und Laptenko O (2005) Bacterial transcription elongation factors: new insights into molecular mechanism of action. *Mol Microbiol* **55:** 1315-1324

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

Bown JA, Owens JT, Meares CF, Fujita N, Ishihama A, Busby SJ und Minchin SD (1999) Organization of open complexes at *Escherichia coli* promoters. Location of promoter DNA sites close to region 2.5 of the sigma70 subunit of RNA polymerase. *J Biol Chem* **274:** 2263-2270

Brendel V, Hamm GH und Trifonov EN (1986) Terminators of transcription with RNA polymerase from *Escherichia coli*: what they look like and how to find them. *J Biomol Struct Dyn* **3:** 705-723

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

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

Burmann BM, Knauer SH, Sevostyana A, Schweimer K, Mooney RA, Landick R, Artsimovitch I und Rösch P (2012) An α-helix to β-barrel domain switch transforms the transcription factor RfaH into a translation factor. *Cell* **150:** 291-303

Burmann BM, Scheckenhofer U, Schweimer K und Rösch P (2011) Domain interactions of the transcription-translation coupling factor *Escherichia coli* NusG are intermolecular and transient. *Biochem J* **435:** 783-789

Burmann BM, Schweimer K, Luo X, Wahl MC, Stitt BL, Gottesman ME und Rösch P (2010) A NusE:NusG Complex Links Transcription and Translation. *Science* **328:** 501-504

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:** 307-16

Canals A, Uson I und Coll M (2010) The structure of RNA-free Rho termination factor indicates a dynamic mechanism of transcript capture. *J Mol Biol* **400:** 16-23

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

Carpousis AJ, Gralla JD (1980) Cycling of ribonucleic acid polymerase to produce oligonucleotides during initiation *in vitro* at the lac UV5 promoter. *Biochemistry* **19:** 3245-3253

Cheeran A, Kolli NR und Sen R (2007) The site of action of the antiterminator protein N from the lambdoid phage H-19B. *J Biol Chem* **282:** 30997-31007

Cheng SW, Court DL und Friedman DI (1995) Transcription termination signals in the nin region of bacteriophage lambda: identification of Rho-dependent termination regions. *Genetics* **140:** 875-887

Ciampi MS (2006) Rho-dependent terminators and transcription termination. *Microbiology* **152:** 2515-2528

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

Costantino N, Zuber M und Court D (1990) Analysis of mutations in the ninR region of bacteriophage lambda that bypass a requirement for lambda N antitermination. *J Bacteriol* **172:** 4610-5

Cramer P (2007) Gene transcription: extending the message. *Nature* **448:** 142-143

Cramer P, Bushnell DA, Fu J, Gnatt AL, Maier-Davis B, Thompson NE, Burgess RR, Edwards AM, David PR und Kornberg RD (2000) Architecture of RNA polymerase II and implications for the transcription mechanism. *Science* **288:** 640-649

Cramer P, Bushnell DA und Kornberg RD (2001) Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. *Science* **292:** 1863-1876

Czyz A, Mooney RA, Iaconi A und Landick R (2014) Mycobacterial RNA polymerase requires a U-tract at intrinsic terminators and is aided by NusG at suboptimal terminators. *MBio* **5:** e00931-14

Das A (1993) Control of transcription termination by RNA-binding proteins. *Annu Rev Biochem* **62:** 893-930

de Vries SJ, van Dijk M und Bonvin AM (2010) The HADDOCK web server for data-driven biomolecular docking. *Nat Protoc* **5:** 883-897

deHaseth PL, Helmann JD (1995) Open complex formation by *Escherichia coli* RNA polymerase: the mechanism of polymerase-induced strand separation of double helical DNA. *Mol Microbiol* **16:** 817-824

deHaseth PL, Zupancic ML und Record MT,Jr (1998) RNA polymerase-promoter interactions: the comings and goings of RNA polymerase. *J Bacteriol* **180:** 3019-3025

Downing WL, Sullivan SL, Gottesman ME und Dennis PP (1990) Sequence and transcriptional pattern of the essential *Escherichia coli* secE-nusG operon. *J Bacteriol* **172:** 1621-1627

Drögemüller J, Stegmann CM, Mandal A, Steiner T, Burmann BM, Gottesman ME, Wöhrl BM, Rösch P, Wahl MC und Schweimer K (2013) An auto-inhibited state in the crystal structure of *Thermotoga maritima* NusG. *Structure* **21:** 365-75

Ebright RH (2000) RNA Polymerase: Structural Similarities Between Bacterial RNA Polymerase and Eukaryotic RNA Polymerase II. *JMB* **304:** 687-698

Ebright RH, Busby S (1995) The *Escherichia coli* RNA polymerase alpha subunit: structure and function. *Curr Opin Genet Dev* **5:** 197-203

Eisenmann A, Schwarz S, Prasch S, Schweimer K und Rösch P (2005) The *E. coli* NusA carboxy-terminal domains are structurally similar and show specific RNAP- and lambdaN interaction. *Protein Sci* **14:** 2018-29

Eisenmann A, Schwarz S, Rösch P und Schweimer K (2004) Sequence-specific ^{1}H, ^{13}C, ^{15}N resonance assignments and secondary structure of the carboxyterminal domain of the *E. coli* transcription factor NusA. *J Biomol NMR* **28:** 193-4

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, Dutta D, Wade J und Nudler E (2010) An allosteric mechanism of Rho-dependent transcription termination. *Nature* **463:** 245-249

Farnham PJ, Greenblatt J und Platt T (1982) Effects of NusA protein on transcription termination in the tryptophan operon of *Escherichia coli*. *Cell* **29:** 945-951

Feig M, Burton ZF (2010) RNA polymerase II with open and closed trigger loops: active site dynamics and nucleic acid translocation. *Biophys J* **99:** 2577-2586

Fiaux J, Bertelsen EB, Horwich AL und Wuthrich K (2002) NMR analysis of a 900K GroEL-GroES complex. *Nature* **418:** 207-211

Fouqueau T, Zeller ME, Cheung AC, Cramer P und Thomm M (2013) The RNA polymerase trigger loop functions in all three phases of the transcription cycle. *Nucleic Acids Res* **41:** 7048-7059

Friedman DI, Schauer AT, Baumann MR, Baron S und Adyha SL (1981) Evidence that ribosomal protein S10 participates in control of transcription termination. *Proc Natl Acad Sci* **78:** 1115-1118

Friedman DI, Olson ER, Georgopoulos C, Tilly K, Herskowitz I und Banuett F (1984) Interactions of bacteriophage and host macromolecules in the growth of bacteriophage lambda. *Microbiol Rev* **48:** 299-325

Gardner KH, Rosen MK und Kay LE (1997) Global folds of highly deuterated, methyl-protonated proteins by multidimensional NMR. *Biochemistry* **36:** 1389-1401

Gaspari M, Larsson NG und Gustafsson CM (2004) The transcription machinery in mammalian mitochondria. *Biochim Biophys Acta* **1659:** 148-152

Geiselmann J, Wang Y, Seifried SE und von Hippel PH (1993) A physical model for the translocation and helicase activities of *Escherichia coli* transcription termination protein Rho. *Proc Natl Acad Sci* **90:** 7754-7758

Gentry D, Xiao H, Burgess R und Cashel M (1991) The omega subunit of *Escherichia coli* K-12 RNA polymerase is not required for stringent RNA control *in vivo*. *J Bacteriol* **173:** 3901-3903

Ghosh P, Ishihama A und Chatterji D (2001) *Escherichia coli* RNA polymerase subunit omega and its N-terminal domain bind full-length beta' to facilitate incorporation into the alpha$_2$beta subassembly. *Eur J Biochem* **268:** 4621-4627

Gotta SL, Miller OL,Jr und French SL (1991) rRNA transcription rate in *Escherichia coli*. *J Bacteriol* **173:** 6647-6649

Greenblatt J, Nodwell JR und Mason SW (1993) Transcriptional antitermination. *Nature* **364:** 401-406

Greenfield N, Fasman GD (1969) Computed circular dichroism spectra for the evaluation of protein conformation. *Biochemistry* **8:** 4108-4116

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

Guo Y, Lew CM und Gralla JD (2000) Promoter opening by sigma(54) and sigma(70) RNA polymerases: sigma factor-directed alterations in the mechanism and tightness of control. *Genes Dev* **14:** 2242-2255

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

Ha KS, Toulokhonov I, Vassylyev DG und Landick R (2010) The NusA N-terminal domain is necessary and sufficient for enhancement of transcriptional pausing via interaction with the RNA exit channel of RNA polymerase. *J Mol Biol* **401:** 708-25

Hartmann G, Honikel KO, Knusel F und Nuesch J (1967) The specific inhibition of the DNA-directed RNA synthesis by rifamycin. *Biochim Biophys Acta* **145:** 843-844

Haugen SP, Ross W und Gourse RL (2008) Advances in bacterial promoter recognition and its control by factors that do not bind DNA. *Nat Rev Microbiol* **6:** 507-519

Heil A, Zillig W (1970) Reconstitution of bacterial DNA-dependent RNA-polymerase from isolated subunits as a tool for the elucidation of the role of the subunits in transcription. *FEBS Lett* **11:** 165-168

Holmes SF, Santangelo TJ, Cunningham CK, Roberts JW und Erie DA (2006) Kinetic investigation of *Escherichia coli* RNA polymerase mutants that influence nucleotide discrimination and transcription fidelity. *J Biol Chem* **281:** 18677-18683

Huber R, Langworthy TA, König H, Thomm M, Wöse CR, Sleytr UB und Stetter KO (1986) *Thermotoga maritima* sp. nov. represents a new genus of unique extremely thermophilic euhacteria growing up to 90 °C. *Arch Microbiol* **144:** 324-333

Irschik H, Gerth K, Hofle G, Kohl W und Reichenbach H (1983) The myxopyronins, new inhibitors of bacterial RNA synthesis from *Myxococcus fulvus* (*Myxobacterales*). *J Antibiot* **36:** 1651-1658

Ishihama A (1992) Role of the RNA polymerase alpha subunit in transcription activation. *Mol Microbiol* **6:** 3283-3288

Ishihama A (1981) Subunit of assembly of *Escherichia coli* RNA polymerase. *Adv Biophys* **14:** 1-35

Jeon YH, Negishi T, Shirakawa M, Yamazaki T, Fujita N, Ishihama A und Kyogoku Y (1995) Solution structure of the activator contact domain of the RNA polymerase alpha subunit. *Science* **270:** 1495-1497

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

Jin DJ, Burgess RR, Richardson JP und Gross CA (1992) Termination efficiency at rho-dependent terminators depends on kinetic coupling between RNA polymerase and rho. *Proc Natl Acad Sci* **89:** 1453-1457

Jones DT (1999) Protein secondary structure prediction based on position-specific scoring matrices. *J Mol Biol* **292:** 195-202

Kalyani BS, Kunamneni R, Wal M, Ranjan A und Sen R (2014) A NusG paralogue from *Mycobacterium tuberculosis*, Rv0639, has evolved to interact with ribosomal protein S10 (Rv0700) but not to function as a transcription elongation-termination factor. *Microbiology* **161:** 67-83

Kalyani BS, Muteeb G, Qayyum MZ und Sen R (2011) Interaction with the nascent RNA is a prerequisite for the recruitment of Rho to the transcription elongation complex *in vitro*. *J Mol Biol* **413:** 548-560

Kapanidis AN, Margeat E, Ho SO, Kortkhonjia E, Weiss S und Ebright RH (2006) Initial transcription by RNA polymerase proceeds through a DNA-scrunching mechanism. *Science* **314:** 1144-1147

Kaplan CD, Larsson KM und Kornberg RD (2008) The RNA polymerase II trigger loop functions in substrate selection and is directly targeted by alpha-amanitin. *Mol Cell* **30:** 547-556

Kassavetis GA, Chamberlin MJ (1981) Pausing and termination of transcription within the early region of bacteriophage T7 DNA *in vitro*. *J Biol Chem* **256:** 2777-2786

Kelley LA, Sternberg MJ (2009) Protein structure prediction on the Web: a case study using the Phyre server. *Nat Protoc* **4:** 363-371

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, Nedialkov YA, Cremona GH, Purtov YA, Lubkowska L, Malagon F, Burton ZF, Strathern JN und Kashlev M (2008) Transient reversal of RNA polymerase II active site closing controls fidelity of transcription elongation. *Mol Cell* **30:** 557-566

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

Koslover DJ, Fazal FM, Mooney RA, Landick R und Block SM (2012) Binding and translocation of termination factor rho studied at the single-molecule level. *J Mol Biol* **423:** 664-676

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

Lau LF, Roberts JW und Wu R (1983) RNA polymerase pausing and transcript release at the lambda tR1 terminator *in vitro*. *J Biol Chem* **258:** 9391-9397

Lee DJ, Minchin SD und Busby SJ (2012) Activating transcription in bacteria. *Annu Rev Microbiol* **66:** 125-152

Leela JK, Syeda AH, Anupama K und Gowrishankar J (2013) Rho-dependent transcription termination is essential to prevent excessive genome-wide R-loops in *Escherichia coli*. *Proc Natl Acad Sci* **110:** 258-263

Linn T, Greenblatt J (1992) The NusA and NusG proteins of *Escherichia coli* increase the *in vitro* readthrough frequency of a transcriptional attenuator preceding the gene for the beta subunit of RNA polymerase. *J Biol Chem* **267:** 1449-1454

Liu X, Bushnell DA, Silva DA, Huang X und Kornberg RD (2011) Initiation complex structure and promoter proofreading. *Science* **333:** 633-637

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

Luttgen H, Robelek R, Muhlberger 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. *JMB* **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

Martinez-Rucobo FW, Sainsbury S, Cheung AC und Cramer P (2011) Architecture of the RNA polymerase-Spt4/5 complex and basis of universal transcription processivity. *EMBO J* **30:** 1302-1310

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 (1992a) Direct interaction between two *Escherichia coli* transcription antitermination factors, NusB and ribosomal protein S10. *J Mol Biol* **223:** 55-66

Mason SW, Li J und Greenblatt J (1992b) Host factor requirements for processive antitermination of transcription and suppression of pausing by the N protein of bacteriophage lambda. *J Biol Chem* **267:** 19418-26

Mishra S, Mohan S, Godavarthi S und Sen R (2013) The interaction surface of a bacterial transcription elongation factor required for complex formation with an antiterminator during transcription antitermination. *J Biol Chem* **288:** 28089-28103

Miwa Y, Horiguchi T und Shigesada K (1995) Structural and functional dissections of transcription termination factor rho by random mutagenesis. *J Mol Biol* **254:** 815-837

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

Mogridge J, Mah TF und Greenblatt J (1995) A protein-RNA interaction network facilitates the template-independent cooperative assembly on RNA polymerase of a stable antitermination complex containing the lambda N protein. *Genes Dev* **9:** 2831-2845

Molodtsov V, Nawarathne IN, Scharf NT, Kirchhoff PD, Showalter HD, Garcia GA und Murakami KS (2013) X-ray crystal structures of the *Escherichia coli* RNA polymerase in complex with benzoxazinorifamycins. *J Med Chem* **56:** 4758-4763

Mooney RA, Artsimovitch I und Landick R (1998) Information processing by RNA polymerase: recognition of regulatory signals during RNA chain elongation. *J Bacteriol* **180:** 3265-3275

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

Morgan WD, Bear DG, Litchman BL und von Hippel PH (1985) RNA sequence and secondary structure requirements for rho-dependent transcription termination. *Nucleic Acids Res* **13:** 3739-3754

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

Mukhopadhyay J, Kapanidis AN, Mekler V, Kortkhonjia E, Ebright YW und Ebright RH (2001) Translocation of sigma(70) with RNA polymerase during transcription: fluorescence resonance energy transfer assay for movement relative to DNA. *Cell* **106:** 453-463

Murakami KS (2013) X-ray crystal structure of *Escherichia coli* RNA polymerase sigma70 holoenzyme. *J Biol Chem* **288:** 9126-9134

Murakami KS, Masuda S und Darst SA (2002) Structural basis of transcription initiation: RNA polymerase holoenzyme at 4 Å resolution. *Science* **296:** 1280-1284

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* **94:** 6641-6645

Muteeb G, Dey D, Mishra S und Sen R (2012) A multipronged strategy of an anti-terminator protein to overcome Rho-dependent transcription termination. *Nucleic Acids Res* **40:** 11213-11228

Nicholson LK, Kay LE, Baldisseri DM, Arango J, Young PE, Bax A und Torchia DA (1992) Dynamics of methyl groups in proteins as studied by proton-detected ^{13}C NMR spectroscopy. Application to the leucine residues of staphylococcal nuclease. *Biochemistry* **31:** 5253-5263

Oppenheim AB, Kobiler O, Stavans J, Court DL und Adhya S (2005) Switches in bacteriophage lambda development. *Annu Rev Genet* **39:** 409-429

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

Palm P, Heil A, Boyd D, Grampp B und Zillig W (1975) The Reconstitution of *Escherichia coli* DNA-Dependent RNA Polymerase from Its Isolated Subunits. *Eur J Biochem* **53:** 283-291

Park JS, Roberts JW (2006) Role of DNA bubble rewinding in enzymatic transcription termination. *Proc Natl Acad Sci* **103:** 4870-4875

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

Peters JM, Mooney RA, Kuan PF, Rowland JL, Keles S und Landick R (2009) Rho directs widespread termination of intragenc and stable RNA transcription. *Proc Natl Acad Soc* **106:** 15406-15411

Platt T (1986) Transcription termination and the regulation of gene expression. *Annu Rev Biochem* **55:** 339-372

Plevin M, Boisbouvier J (2012) Isotope-labelling of methyl groups for NMR studies of large proteins. In M Clore, J Potts, eds, Recent Developments in Biomolecular NMR. The Royal Society of Chemistry, pp FP001-347

Prasch S, Jurk M, Washburn RS, Gottesman ME, Wöhrl BM und Rösch P (2009) RNA-binding specificity of *E. coli* NusA. *Nucl Acids Res* **37:** 4736-4742

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

Ralling G, Linn T (1987) Evidence that Rho and NusA are involved in termination in the rplL-rpoB intercistronic region. *J Bacteriol* **169:** 2277-2280

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

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* **93:** 342-346

Revyakin A, Liu C, Ebright RH und Strick TR (2006) Abortive initiation and productive initiation by RNA polymerase involve DNA scrunching. *Science* **314:** 1139-1143

Richardson JP (2002) Rho-dependent termination and ATPases in transcript termination. *Biochim Biophys Acta* **1577:** 251-260

Richardson JP, Grimley C und Lowery C (1975) Transcription termination factor rho activity is altered in *Escherichia coli* with *suA* gene mutations. *Proc Natl Acad Sci* **72:** 1725-1728

Richardson LV, Richardson JP (1996) Rho-dependent termination of transcription is governed primarily by the upstream Rho utilization (*rut*) sequences of a terminator. *J Biol Chem* **271:** 21597-21603

Roberts JW, Shankar S und Filter JJ (2008) RNA polymerase elongation factors. *Annu Rev Microbiol* **62:** 211-233

Rosen MK, Gardner KH, Willis RC, Parris WE, Pawson T und Kay LE (1996) Selective methyl group protonation of perdeuterated proteins. *J Mol Biol* **263:** 627-636

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

Roy A, Kucukural A und Zhang Y (2010) I-TASSER: a unified platform for automated protein structure and function prediction. *Nat Protoc* **5:** 725-738

Ruschak AM, Kay LE (2010) Methyl groups as probes of supra-molecular structure, dynamics and function. *J Biomol NMR* **46:** 75-87

Sakata-Sogawa K, Shimamoto N (2004) RNA polymerase can track a DNA groove during promoter search. *Proc Natl Acad Sci* **101:** 14731-14735

Salstrom JS, Szybalski W (1978) *Coli* phage lambda nutL-: a unique class of mutants defective in the site of gene N product utilization for antitermination of leftward transcription. *J Mol Biol* **124:** 195-221

Salzmann M, Pervushin K, Wider G, Senn H und Wüthrich K (2000) NMR Assignment and Secondary Structure Determination of an Octameric 110 kDa Protein Using TROSY in Triple Resonance Experiments. *J Am Chem Soc* **122:** 7543-7548

Santangelo TJ, Artsimovitch I (2011) Termination and antitermination: RNA polymerase runs a stop sign. *Nat Rev Microbiol* **9:** 319-329

Santangelo TJ, Roberts JW (2004) Forward translocation is the natural pathway of RNA release at an intrinsic terminator. *Mol Cell* **14:** 117-126

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

Schweimer K, Prasch S, Santhanam SP, Bubunenko M, Gottesman ME und Rösch P (2011) NusA interaction with the α-subunit of *E. coli* RNA Polymerase is *via* the UP-element site and releases autoinhibition. *Structure* **19:** 945-954

Sekine S, Tagami S und Yokoyama S (2012) Structural basis of transcription by bacterial and eukaryotic RNA polymerases. *Curr Opin Struct Biol* **22:** 110-118

Sevostyanova A, Artsimovitch I (2010) Functional analysis of *Thermus thermophilus* transcription factor NusG. *Nucleic Acids Res* **38:** 7432-7445

Sevostyanova A, Belogurov GA, Mooney RA, Landick R und Artsimovitch I (2011) The β subunit gate loop is required for RNA polymerase modification by RfaH and NusG. *Mol Cell* **43:** 253-262

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

Shin DH, Nguyen HH, Jancarik J, Yokota H, Kim R und Kim SH (2003) Crystal structure of NusA from *Thermotoga maritima* and functional implication of the N-terminal domain. *Biochemistry* **42:** 13429-13437

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, Sosunova E, Mustaev A, Bass I, Nikiforov V und Goldfarb A (2003) Unified two-metal mechanism of RNA synthesis and degradation by RNA polymerase. *EMBO J* **22:** 2234-2244

Sprangers R, Kay LE (2007) Quantitative dynamics and binding studies of the 20S proteasome by NMR. *Nature* **445:** 618-622

Stagno JR, Altieri AS, Bubunenko M, Tarasov SG, Li J, Court DL, Byrd RA und Ji X (2011) Structural basis for RNA recognition by NusB and NusE in the initiation of transcription antitermination. *Nucleic Acids Res* **39:** 7803-7815

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

Steinmetz EJ, Platt T (1994) Evidence supporting a tethered tracking model for helicase activity of *Escherichia coli* Rho factor. *Proc Natl Acad Sci* **91:** 1401-1405

Storz G, Opdyke JA und Wassarman KM (2006) Regulating bacterial transcription with small RNAs. *Cold Spring Harb Symp Quant Biol* **71:** 269-273

Suh WC, Leirmo S und Record MT,Jr (1992) Roles of Mg^{2+} in the mechanism of formation and dissociation of open complexes between *Escherichia coli* RNA polymerase and the lambda PR promoter: kinetic evidence for a second open complex requiring Mg^{2+}. *Biochemistry* **31:** 7815-7825

Sullivan SL, Gottesman ME (1992) Requirement for *E. coli* NusG protein in factor-dependent transcription termination. *Cell* **68:** 989-994

Tagami S, Sekine S, Kumarevel T, Hino N, Murayama Y, Kamegamori S, Yamamoto M, Sakamoto K und Yokoyama S (2010) Crystal structure of bacterial RNA polymerase bound with a transcription inhibitor protein. *Nature* **468:** 978-982

Tan R, Frankel AD (1995) Structural variety of arginine-rich RNA-binding peptides. *Proc Natl Acad Sci* **92:** 5282-5286

Tang H, Severinov K, Goldfarb A und Ebright RH (1995) Rapid RNA polymerase genetics: one-day, no-column preparation of reconstituted recombinant *Escherichia coli* RNA polymerase. *Proc Natl Acad Sci* **92:** 4902-4906

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

Thomsen ND, Berger JM (2009) Running in reverse: the structural basis for translocation polarity in hexameric helicases. *Cell* **139:** 523-534

Tomar SK, Artsimovitch I (2013) NusG-Spt5 Proteins-Universal Tools for Transcription Modification and Communication. *Chem Rev* **131:** 8604-19

Tomar SK, Knauer SH, Nandymazumdar M, Rösch P und Artsimovitch I (2013) Interdomain contacts control folding of transcription factor RfaH. *Nucleic Acids Res* **41:** 10077-85

Toulokhonov I, Artsimovitch I und Landick R (2001) Allosteric control of RNA polymerase by a site that contacts nascent RNA hairpins. *Science* **292:** 730-733

Toulokhonov 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

Tugarinov V, Hwang PM, Ollerenshaw JE und Kay LE (2003) Cross-correlated relaxation enhanced ^{1}H-^{13}C NMR spectroscopy of methyl groups in very high molecular weight proteins and protein complexes. *J Am Chem Soc* **125:** 10420-10428

Tugarinov V, Muhandiram R, Ayed A und Kay LE (2002) Four-dimensional NMR spectroscopy of a 723-residue protein: chemical shift assignments and secondary structure of malate synthase g. *J Am Chem Soc* **124:** 10025-10035

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 A 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

Vo NV, Hsu LM, Kane CM und Chamberlin MJ (2003) *In vitro* studies of transcript initiation by *Escherichia coli* RNA polymerase. 3. Influences of individual DNA elements within the promoter recognition region on abortive initiation and promoter escape. *Biochemistry* **42:** 3798-3811

Vogel U, Jensen KF (1994) The RNA chain elongation rate in *Escherichia coli* depends on the growth rate. *J Bacteriol* **176:** 2807-2813

Walstrom KM, Dozono JM und von Hippel PH (1997) Kinetics of the RNA-DNA helicase activity of *Escherichia coli* transcription termination factor rho. 2. Processivity, ATP consumption, and RNA binding. *Biochemistry* **36:** 7993-8004

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 Y, Severinov K, Loizos N, Fenyo D, Heyduk E, Heyduk T, Chait BT und Darst SA (1997) Determinants for *Escherichia coli* RNA polymerase assembly within the beta subunit. *J Mol Biol* **270:** 648-662

Wei RR, Richardson JP (2001) Mutational changes of conserved residues in the Q-loop region of transcription factor Rho greatly reduce secondary site RNA-binding. *J Mol Biol* **314:** 1007-1015

Weisberg RA (2008) Transcription by Moonlight: Structural Basis of an Extraribosomal Activity of Ribosomal Protein S10. *Mol Cell* **32:** 747-748

Weixlbaumer A, Leon K, Landick R und Darst SA (2013) Structural basis of transcriptional pausing in bacteria. *Cell* **152:** 431-441

Werner F (2012) A nexus for gene expression-molecular mechanisms of Spt5 and NusG in the three domains of life. *J Mol Biol* **417:** 13-27

Werner F, Grohmann D (2011) Evolution of multisubunit RNA polymerases in the three domains of life. *Nat Rev Microbiol* **9:** 85-98

Whalen WA, Das A (1990) Action of an RNA site at a distance: role of the *nut* genetic signal in transcription antitermination by phage-lambda N gene product. *New Biol* **2:** 975-991

Wimberly BT, Brodersen DE, Clemons WMJ, Morgan-Warren RJ, Carter AP, Vonrhein C, Hartsch T und Ramakrishnan V (2000) Structure of the 30S ribosomal subunit. *Nature* **407:** 327-339

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

Yang X, Lewis PJ (2010) The interaction between RNA polymerase and the elongation factor NusA. *RNA Biol* **7:** 272-275

Yang X, Molimau S, Doherty GP, Johnston EB, Marles-Wright J, Rothnagel R, Hankamer B, Lewis RJ und Lewis PJ (2009) The structure of bacterial RNA polymerase in complex with the essential transcription elongation factor NusA. *EMBO Rep* **10:** 997-1002

Yarnell WS, Roberts JW (1999) Mechanism of intrinsic transcription termination and antitermination. *Science* **284:** 611-615

Yee D, Armstrong VW und Eckstein F (1979) Mechanistic studies on deoxyribonucleic acid dependent ribonucleic acid polymerase from *Escherichia coli* using phosphorothioate analogues. 1. Initiation and pyrophosphate exchange reactions. *Biochemistry* **18:** 4116-4120

Zaychikov E, Denissova L, Meier T, Gotte M und Heumann H (1997) Influence of Mg^{2+} and temperature on formation of the transcription bubble. *J Biol Chem* **272:** 2259-2267

Zaychikov E, Martin E, Denissova L, Kozlov M, Markovtsov V, Kashlev M, Heumann H, Nikiforov V, Goldfarb A und Mustaev A (1996) Mapping of catalytic residues in the RNA polymerase active center. *Science* **273:** 107-109

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 A resolution. *Cell* **98:** 811-824

Zhang J, Landick R (2009) Chapter 7 Substrate Loading, Nucleotide Addition, and Translocation by RNA Polymerase. RNA Polymerases as Molecular Motors. *The Royal Society of Chemistry*, pp 206-235

Zhu AQ, von Hippel PH (1998) Rho-dependent termination within the trp t' terminator. II. Effects of kinetic competition and rho processivity. *Biochemistry* **37:** 11215-11222

6. Publikationsliste

6.1 Einzelarbeit A

Martin Strauß, Kristian Schweimer, Björn M. Burmann, Anne Richter, Stephanie Güttler, Birgitta M. Wöhrl, Paul Rösch (2015): **The two domains of *Mycobacterium tuberculosis* NusG protein are dynamically independent**, *Journal of Biomolecular Structure and Dynamics*, akzeptiert am 16.03.2015.

Die Klonierung des Gens für *Mt*NusG, die Proteinreinigung und die Aufnahme der entsprechenden NMR-Spektren wurden von Björn M. Burmann und Stephanie Güttler durchgeführt. Die Gene für *Mt*NusG-CTD und *Mt*NusG-NTD wurden, mit Hilfe von Birgitta M. Wöhrl, von mir kloniert. Die Proteinreinigung, die Aufnahme der NMR-Spektren, die Relaxationsmessungen und die Strukturbestimmung der *Mt*NusG-CTD wurden von mir und Kristian Schweimer durchgeführt. Ich habe Anne Richter bei der Reinigung und Analyse der *Mt*NusG-NTD betreut. Die Analyse der Ergebnisse erfolgte durch mich in Absprache mit Paul Rösch und Kristian Schweimer. Die Strukturmodelle der *Mt*NusG-NTD mittels I-TASSER wurden von mir berechnet. Das Manuskript wurde von Björn M. Burmann, Paul Rösch, Kristian Schweimer und mir verfasst.

6.2 Einzelarbeit B

Martin Strauß, Christal Vitiello, Kristian Schweimer, Max E. Gottesman, Paul Rösch, Stefan H. Knauer (2016): **Transcription is regulated by NusA:NusG interaction**, *Nucleic Acids Research,* Mai 2016.

Die Proteinreinigungen wurden von mir durchgeführt. Die NMR-Daten wurden von mir und Kristian Schweimer aufgenommen, ausgewertet und analysiert. Die Ergebnisse wurden mit Paul Rösch, Johanna Drögemüller und Stefan H. Knauer diskutiert. Der *in vitro* Transkriptionsassay wurde von Christal Vitiello durchgeführt, die *in vivo* Messungen von Max E. Gottesman. Das Manuskript wurde von Paul Rösch, Max E. Gottesman, Stefan H. Knauer und mir verfasst.

6.3 Einzelarbeit C

Johanna Drögemüller*, Martin Strauß*, Kristian Schweimer, Birgitta M. Wöhrl, Stefan H. Knauer, Paul Rösch (2015): **Exploring RNA polymerase regulation by NMR Spectroscopy**, *Scientific Reports,* Juni 2015 4;5: 10825.

* Beide Autoren haben in gleichem Maße zur Arbeit beigetragen.

Die Klonierung der Gene *rpoA* und *rpoB*, sowie die Reinigung der α- und β-Untereinheit der RNAP wurden von mir durchgeführt, wobei Birgitta M. Wöhrl bei der Planung der Klonierung beraten hat. Johanna Drögemüller klonierte die Gene *rpoC* und *rpoZ* und reinigte die daraus exprimierten RNAP-Untereinheiten β‘ und ω. Der Zusammenbau und die Reinigung der RNAP wurden von Johanna Drögemüller und mir gemeinsam entwickelt und ausgeführt. Die Reinigung aller weiteren Proteine und die Aufnahme der CD-Spektren erfolgte durch Johanna Drögemüller und mich. Die Bestimmung der mit den Nus-Faktoren interagierenden RNAP-Domäne, sowie die Aufnahme der dazugehörigen Spektren, erfolgten durch Johanna Drögemüller und mich. Kristian Schweimer hat die für die RNAP-Messungen verwendeten NMR-Pulsprogramme vorbereitet und angepasst. Das Manuskript wurde von Paul Rösch, Stefan H. Knauer, Johanna Drögemüller und mir verfasst.

6.4 Einzelarbeit D

Johanna Drögemüller*, Martin Strauß*, Kristian Schweimer, Marcel Jurk, Paul Rösch, Stefan H. Knauer (2015): **Determination of RNA polymerase binding surfaces of transcription factors by NMR spectroscopy**, *Scientific Reports,* November 2015.

* Beide Autoren haben in gleichem Maße zur Arbeit beigetragen.

Die Struktur der NusA-NTD wurde von Marcel Jurk unter Leitung von Kristian Schweimer bestimmt. Die Reinigung der Proteine erfolgte durch Johanna Drögemüller und mich. Die Messmethode zur Bestimmung der Bindungsfläche wurde von Johanna Drögemüller und mir entwickelt. Die Messungen zu NusA-NTD und NusE:NusB wurden von mir durchgeführt, während die Messungen für NusG-NTD von Johanna Drögemüller durchgeführt wurden. Die Auswertung der Ergebnisse für die NusA-NTD erfolgte durch mich, während Johanna Drögemüller die Ergebnisse für NusG-NTD und NusE:NusB auswertete. Die Planung, Durchführung und Auswertung der Verdrängungsexperimente mit RNAP, NusE:NusB und NusG-CTD geschah durch Johanna Drögemüller und mich. Stefan H. Knauer erzeugte die Docking-Modelle mittels HADDOCK (de Vries *et al.*, 2010). Das Manuskript wurde von Paul Rösch, Stefan H. Knauer, Johanna Drögemüller und mir verfasst.

7. Einzelarbeiten

7.1 Einzelarbeit A

Martin Strauß, Kristian Schweimer, Björn M. Burmann, Anne Richter, Stephanie Güttler, Birgitta M. Wöhrl, Paul Rösch (2015): **The two domains of *Mycobacterium tuberculosis* NusG protein are dynamically independent**, *Journal of Biomolecular Structure and Dynamics*, Mai 2015 1:1-10.

Journal of Biomolecular Structure and Dynamics, 2015
http://dx.doi.org/10.1080/07391102.2015.1031700

The two domains of *Mycobacterium tuberculosis* NusG protein are dynamically independent

Martin Strauß[a]*, Kristian Schweimer[a], Björn M. Burmann[a,b], Anne Richter[a], Stephanie Güttler[c], Birgitta M. Wöhrl[a] and Paul Rösch[a]

[a]*Lehrstuhl Biopolymere und Forschungszentrum für Bio-Makromoleküle, Universität Bayreuth, Universitätsstraße 30, 95447 Bayreuth, Germany;* [b]*Biozentrum, University of Basel, Klingelbergstr. 70, 4056 Basel, Switzerland;* [c]*Lehrstuhl Tierökologie II, Universität Bayreuth, Universitätsstraße 30, 95447 Bayreuth, Germany*

Communicated by Ramaswamy H. Sarma

(*Received 22 December 2014; accepted 16 March 2015*)

Transcription elongation factor NusG from *Escherichia coli* couples transcription and translation. It is the only conserved transcription factor in all three kingdoms of life, playing a variety of roles in gene expression. *E. coli* NusG consists of two non-interacting domains. While the N-terminal domain interacts with RNA polymerase, the C-terminal domain contacts NusE (S10), or the Rho transcription termination factor. The two corresponding domains of *Thermotoga maritima* NusG are mutually interacting. Therefore, NusG here forms an autoinhibited state, where the binding sites to RNAP, NusE, and the Rho factor are masked. Recent functional studies showed differences between NusG from *E. coli* and *Mycobacterium tuberculosis*. In contrast to *E. coli* NusG, *M. tuberculosis* NusG is able to stimulate intrinsic termination, but is not able to bind the Rho factor. To analyze whether this has structural reasons, we determined the solution structure of the carboxyterminal domain of *M. tuberculosis* NusG by nuclear magnetic resonance spectroscopy. Furthermore, we modeled the wild-type full-length protein, and present evidence that the two domains of this protein do not interact in solution by NMR dynamics measurements.

Keywords: bacterial transcription; NusG; RfaH; NMR; *Mycobacterium tuberculosis*

Introduction

Transcription of deoxyribonucleic acid (DNA) to ribonucleic acid (RNA) in bacteria is performed by the RNA-polymerase (RNAP) core enzyme. RNAP moves along the DNA template to synthesize the corresponding RNA strand. This RNAP movement along the DNA is not uniform, but is modulated by nucleic acid sequences, small molecule regulators, and proteins such as the N-utilization substances (Nus) A, B, E, and G (Roberts, Shankar, & Filter, 2008).

NusG is conserved among *archaea* and *bacteria,* and it is highly homologous to the eukaryotic elongation factor Spt5 (Werner, 2012). The *E. coli* NusG amino-terminal domain (NTD) interacts with RNAP and increases the transcription rate of the transcription elongation complex (Mooney, Schweimer, Rösch, Gottesman, & Landick, 2009). The *E. coli* NusG (*Ec*NusG) carboxyterminal domain (CTD) interacts with *E. coli* NusE (ribosomal protein S10) and, alternatively, *E. coli* termination factor Rho (Burmann et al., 2010). The two *Ec*NusG structural domains, CTD and NTD, consisting in total of 181 amino acids, are connected via a flexible linker and move independently (Burmann, Scheckenhofer, Schweimer, & Rösch, 2011; Mooney et al., 2009).

Like *Ec*NusG, NusG proteins from other bacteria such as *Thermus thermophilus* and *Aquifex aeolicus* (*A. aeolicus*) also lack intramolecular domain interactions (Reay et al., 2004; Steiner, Kaiser, Marinkovic, Huber, & Wahl, 2002). Contrasting, the NusG paralogue RfaH of *E. coli* exhibits a NusG-like NTD but a differently folded, α-helical CTD tightly associated with the NTD (Belogurov et al., 2007). Upon domain opening, the RfaH-CTD refolds into an all-β-sheet conformation that is structurally highly similar to NusG-CTD (Burmann et al., 2012). This domain interaction serves as an autoinhibitory mechanism to avoid interference with the general transcription factor NusG. A related closed structure was observed in *Thermotoga maritima* NusG (*Tm*NusG), where the NTD and CTD are mutually interacting within the crystal unit (Drögemüller et al., 2013). This interaction was confirmed by NMR relaxation as well as residual dipolar coupling data in solution, excluding the possibility of crystal packing effects (Drögemüller et al., 2013). In both RfaH and *Tm*NusG, the CTD masks a hydrophobic surface on the NTD that is interacting with the RNAP to modulate transcription (Belogurov et al., 2007; Drögemüller et al., 2013).

*Corresponding author. Email: martin.strauss@uni-bayreuth.de

*Tm*NusG and *A. aeolicus* NusG (*Aa*NusG) contain an additional domain of approximately 60 residues within the NTD. For *Tm*NusG, the additional domain was suggested to be able to bind nucleic acids (Liao, Lurz, Dobrinski, & Dennis, 1996). *Aa*NusG forms domain-swapped dimers under certain crystallization conditions, but the physiological relevance of this dimer remains unclear, as under different conditions *Aa*NusG also crystallizes as a monomer (Knowlton et al., 2003; Steiner et al., 2002).

As the functional details of *Mycobacterium tuberculosis* (*M. tuberculosis*) NusG (*Mt*NusG) remain poorly understood, we investigated the structural properties of this protein. These studies revealed distinct flexible parts that are not found in related NusG proteins studied earlier. Further, we report the high-resolution solution structure of *Mt*NusG-CTD, revealing a typical NusG-CTD-fold, and experimental evidence that NTD and CTD do not strongly mutually interact in this protein.

Materials and methods

Cloning

The *Mt*NusG gene optimized to *E. coli* codon usage was purchased from GenScript (USA) and subsequently cloned into pET11a (Novagen, Germany) via *Nde*I and *BamH*I (pET11a_*Mt*NusG). A DNA fragment coding *Mt*NusG-CTD (residues 178-238) was cloned into pET101/D-TOPO (Invitrogen, Germany) containing an amino-terminal hexa-histidine tag as well as a tobacco etch virus (TEV) protease cleavage site between the His-tag and NusG-CTD (pET101/D-TOPO_6His-*Mt*NusG-CTD). The *Mt*NusG-NTD (residues 1-178) gene was cloned into pET-GB1a (*G. Stier, EMBL,* Heidelberg, Germany) via *Nco*I and *BamH*I, with a GB1-(*streptococcal* immunoglobulin-binding domain of protein G) fusion tag, an amino-terminal hexa-histidine tag, and a TEV protease cleavage site between GB1 and NusG-NTD (pETGB_1a_*Mt*NusG-NTD).

Media

Bacterial cells were grown in 2 l Luria broth (LB; Sambrook, Fritsch, & Maniatis, 1994) in the presence of the required antibiotic (30 µg/ml kanamycin or 100 µg/ml ampicillin). For ^{15}N-labeling, bacteria were grown in M9 minimal medium supplemented with 1 × MEM vitamin solution (Gibco, Germany), TS2 solution, and $(^{15}NH_4)_2SO_4$ (Campro Scientific, Germany) as the only nitrogen source (Meyer & Schlegel, 1983; Sambrook et al., 1994). For ^{15}N/^{13}C double labeled proteins, 4 g/l D-[^{13}C]-glucose (Spectra Stable Isotopes, USA) was used as sole carbon source.

Expression and purification of MtNusG

E. coli BL21 DE3 (NEB, Germany) harboring pET11a_*Mt*NusG was grown at 37 °C to an OD_{600} of 0.6–0.8. Gene expression was induced by adding 0.5 mM isopropyl 1-thio-β-D-galactopyranoside (IPTG). The cells were harvested 3 h after induction by centrifugation (9000 × *g*, 15 min, 4 °C), and the cell pellet was resuspended in 50 mM Tris/HCl, 150 mM NaCl, pH 7.5 (buffer A) four times their weight, and stored at −20 °C overnight. Subsequently, the cells were thawed, and after addition of half a tablet of EDTA-free protease inhibitor (Roche, Germany), they were lysed using a microfluidizer (MFTI Corporation, USA). After centrifugation of the lysate for 30 min at 4 °C, 12,000 g to remove cell debris, 0.313 g/ml ammonium sulfate were added to the supernatant under continuous stirring for 15 min at room temperature and 30 min at 4 °C. Subsequently, it was centrifuged for 20 min at 4000 g, 4 °C. The pellet was dissolved in 40 ml of 20 mM Na-phosphate, pH 7.0 (buffer B), and dialyzed against this buffer overnight. The dialysate was applied to a HiTrap Blue HP column (GE Healthcare, Germany) and eluted with an NaCl step gradient up to 2 M. Fractions containing *Mt*NusG were pooled, dialyzed overnight against buffer B at 4 °C, applied to a HiTrap Heparin HP column (GE Healthcare, Germany), and eluted with a NaCl step gradient up to 2 M. *Mt*NusG containing fractions were dialyzed overnight against buffer B with 150 mM NaCl at 4 °C. The final purification step was performed with a Superdex75 size exclusion column (GE Healthcare, Germany) with buffer B supplemented with 150 mM NaCl. Fractions containing pure *Mt*NusG were pooled, concentrated by ultrafiltration (MWCO = 5000 Da, Sartorius, Germany), and stored at −80 °C after shock-freezing with liquid nitrogen.

Expression and purification of MtNusG-CTD (178-238)

E. coli BL21 DE3 harboring pET101/DTOPO_6His-*Mt*NusG-CTD was grown and harvested as described above. However, induction was performed with 1 mM IPTG, and cells were harvested 4 h after induction. The cell pellet was resuspended in buffer (buffer A + 10 mM imidazole) four times the cell mass and stored at −20 °C. After cell lysis, the supernatant was applied to a 5 ml Ni^{2+}-HisTrap column (GE Healthcare, Germany), and the protein was purified by an imidazole step gradient (0.01–1 M). The fractions containing 6His-*Mt*NusG-CTD were pooled and the His-tag was cleaved off by TEV-protease during dialysis against buffer A at room temperature overnight. The His-tag and the TEV protease were removed by a second Ni^{2+}-affinity chromatography. The untagged *Mt*NusG-CTD protein was collected in the flow-through, concentrated by

ultrafiltration (MWCO = 5000 Da, Sartorius, Germany), and stored at −80 °C after shock-freezing with liquid nitrogen.

Expression and purification of MtNusG-NTD (1-178)

E. coli BL21 DE3 harboring pETGB_1a_*Mt*NusG-NTD (1-178) was grown and harvested as described for the full-length *Mt*NusG. Resuspension, lysis, and purification were as described for *Mt*NusG-CTD.

NMR spectroscopy

NMR spectra were recorded on Bruker Avance 600 MHz, Avance 700 MHz, and Avance 800 MHz spectrometers; the latter two equipped with cryogenically cooled triple resonance probeheads at 298 K. For the assignment of *Mt*NusG-CTD, triple resonance experiments with selective proton pulses (BEST-NMR-experiments) were performed (Lescop, Schanda, & Brutscher, 2007). NMR spectra to assign chemical shifts were obtained from 1 mM [U-^{15}N, ^{13}C]-*Mt*NusG-CTD in 25 mM HEPES, 150 mM NaCl, pH 7.0, and supplemented with 10% D_2O. Three-dimensional (3D) ^{13}C- and ^{15}N-edited NOESY (nuclear Overhauser enhancement spectroscopy) experiments (mixing times 120 ms) were recorded for derivation of distance restraints. Residual dipolar couplings were determined using a sample containing 10 mg/ml Pf1 phages (Hansen, Mueller, & Pardi, 1998; Hyglos, Germany) with in-phase anti-phase methodology (Ottiger, Delaglio, & Bax, 1998). NMR data were processed using in-house software and visualized as well as analyzed with NMRView (Johnson, 2004).

For the characterization of overall and internal motion, ^{15}N longitudinal (R_1) and transverse (R_2) relaxation rates together with $\{^1H\}^{15}N$ steady-state NOEs were recorded using standard methods (Kay, Torchia, & Bax, 1989) at 700.2 MHz ^{1}H frequency at 298 K. Relaxation delays for R_1 and R_2 relaxation experiments were fitted to a mono-exponential decay using the program curve fit (A.G. Palmer, Dept. of Biochemistry and Molecular Biophysics, Columbia University, USA).

Structure calculation

Distances derived from NOEs have been assigned to values of 5, 4, and 3 Å according to their intensities. Dihedral backbone angle restraints were obtained from chemical shifts using the program TALOS (Cornilescu, Delaglio, & Bax, 1999). Structure calculations were performed with XPLOR-NIH 1.2.1 (Schwieters, Kuszewski, Tjandra, & Clore, 2003) using a three-step simulated annealing protocol with floating assignment of prochiral groups including a conformational database potential. The ten structures showing the lowest values of the target function excluding the database potential were further analyzed with X-PLOR-NIH (Schwieters et al., 2003), PyMOL (DeLano, 2007), and PROCHECK 3.5.4 (Laskowski, MacArthur, Moss, & Thornton, 1993). Coordinates were deposited in the Protein Data Bank (PDB accession code: 2MI6), and chemical shift assignments were deposited in the BioMagResBank (BMRB, accession code 19,667).

CD-spectroscopy

Far UV CD measurements were performed on a J-1100 spectropolarimeter with a CDF-426S temperature control unit (JASCO International, Japan). Samples were prepared by dialyzing protein solutions against 10 mM Na-phosphate buffer, pH 7.5. Spectra were recorded at 298 K 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 10 spectra were accumulated. In order to normalize the measured ellipticity, the mean residue molar ellipticity was calculated as:

$$[\theta]_{\mathrm{MRW}} = \theta/(c \cdot d \cdot N) \quad (1)$$

θ is the measured ellipticity; MRW is the mean residue mass; c is the protein concentration; d is the path length; and N is the number of amino acids. The secondary structure content was assessed by using the DichroWeb server with the CDSSTR-algorithm (Sreerama & Woody, 2000).

Computational methods

We used with default parameters: multiple sequence alignment, ClustalW2 (EMBL-EBI; Larkin et al., 2007); graphics of protein structures, PyMOL v1.0 (DeLano, 2007); secondary structure predictions, PSIPRED (Jones, 1999); homology modeling, I-TASSER (Roy, Kucukural, & Zhang, 2010). Calculations of the electrostatic potentials were performed with the PyMOL-included APBS tool. The required pqr-files were prepared using the pdb2pqr web server (AMBER force field; Dolinsky, Nielsen, McCammon, & Baker, 2004).

Results and discussion

Structural analysis of full-length MtNusG

Initial sequence alignments revealed that *Mt*NusG has 40 additional amino-terminal residues compared to *Aa*NusG, *Tm*NusG, and *Ec*NusG (Figure S1), but lacks the domain insertion within the NTD that was found in *Aa*NusG as well as in *Tm*NusG. Furthermore, the CTD:NTD linker of *Mt*NusG is extended in comparison to the other NusG proteins.

The 2D [^{1}H, ^{15}N] heteronuclear single-quantum coherence (HSQC) spectrum of [*U*-^{15}N]-*Mt*NusG (U denotes uniform labeling) displays the typical large dispersion of resonance signals characteristic for a well-folded protein domain (Figure 1(A)). In addition to the well-dispersed signals, several sharp, intense peaks are visible in the random-coil region of the spectrum (Figure 1(B)). Most likely, these signals belong to the residues of the amino-terminal extension and/or the large linker region pointing to the high flexibility of these regions. The large degree of flexibility is additionally indicated by the high degree of disordered elements inferred from the CD-spectra of the full-length protein as well as the individual domains (Table 1, Figure 1(C)). The full-length *Mt*NusG and *Mt*NusG-NTD show flexible parts over 50% of the protein sequence, compared to 40% in the *E. coli* protein (Kalyani, Kunamneni, Wal, Ranjan, & Sen, 2015). However, the stability of the full-length *Mt*NusG protein samples was insufficient to record triple-resonance spectra that would have been necessary for sequence-specific resonance assignment and solution structure determination. Therefore, the two domains were produced, analyzed separately (NTD: residues 1-178; CTD: residues 178-238), and compared with full-length protein subsequently.

Homology models for MtNusG-NTD

The well-dispersed 2D [^{1}H, ^{15}N] HSQC-spectrum of *Mt*NusG-NTD shows nearly perfect overlap with the spectrum of full-length *Mt*NusG, indicative of the native fold for this single-domain construct (Figure 2(A)). As the long-term stability for this construct was insufficient for resonance assignments, we resorted to *in silico* secondary structure prediction by PSIPRED (Jones, 1999), which suggested the N-terminus to be devoid of any

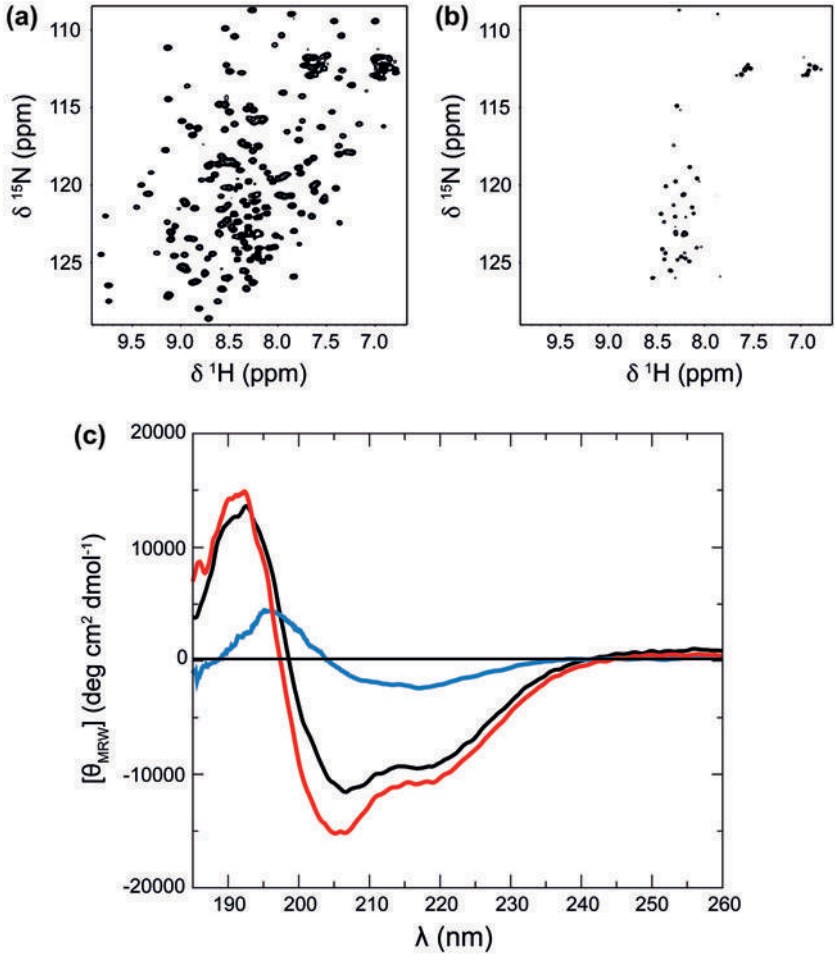


Figure 1. Spectroscopic analysis of full-length *Mt*NusG. (A) [^{1}H, ^{15}N] HSQC spectrum of 0.3 mM [*U*-^{15}N]-*Mt*NusG in 25 mM Na-phosphate, 150 mM NaCl, pH 7.5 at 298 K. (B) Same spectrum as in panel (A) plotted at high contour level, showing the high-intensity resonances. (C) Overlay of the CD-spectra for *Mt*NusG (black), *Mt*NusG-NTD (red), and *Mt*NusG-CTD. Θ_{MRW} vs. wavelength in nm. Θ_{MRW} was calculated according to Equation (1). Temperature, 298 K.

Table 1. Secondary structure content determined by CD-spectroscopy.

Construct	α-helix (%)	β-sheet (%)	β-turns (%)	Unordered (%)
*Mt*NusG	19	27	23	32
*Mt*NusG-NTD	30	19	22	30
*Mt*NusG-CTD	9	38	17	36

stable secondary structure elements up to residue 35 (Figure S2). To obtain a 3D model, we analyzed the *Mt*NusG-NTD sequence by multiple-threading sequence alignment with the I-TASSER webserver (Roy et al., 2010), which calculated five structures (Figure 3). Whereas overall the predictions were almost identical, the relative orientation of the 35–40 amino-terminal residues remained undefined. Based on these *in silico* analyses, we therefore conclude that *Mt*NusG-NTD has a typical NusG-NTD-fold. This fold is highly conserved as NusG-NTD binds to RNAP in all three kingdoms of life (Werner, 2012). Additionally *Mt*NusG-NTD has an additional, flexible, and mainly unstructured amino-terminus. This observation is further supported by the observation that signals within the 2D [^{1}H, ^{15}N] HSQC-spectrum accumulate at chemical shift regions which are characteristic for residues embedded in unstructured parts of proteins ($\delta(^{1}\mathrm{H}) = 7.9$–$8.5$ ppm; Figures 1(B) and 2(A)). The amino-terminus consists mainly of alanines and polar, mostly acidic residues (12 alanine, 4 glutamic acid, 5 aspartic acid, and 4 threonine). The N-terminus was recently proposed to be important for dimer formation by pBlast analysis or for masking a hydrophobic surface on the NTD or CTD on the basis of *in vitro* cross-linking experiments (Kalyani et al., 2015). The high signal intensity of these residues observed in our NMR experiments indicates a high flexibility of this region, essentially ruling out its participation in dimerization and/or binding interface masking. The formation of cross-links might therefore just represent one possible conformation of the amino-terminal amino acids. As the overall charge distributions on the surfaces of *Ec*NusG and *Mt*NusG Δ1-40 are very similar (Figure S3), masking of hydrophobic parts of the latter protein may not be essential in *M. tuberculosis*. Calculations with the ExPASy ProtParam webtool (Gasteiger et al., 2005) showed that the presence of the additional 40 amino acids at the amino-terminus leads to a lowering of the isoelectric point from 8.3 (residues 41-178) to 4.9 (residues 1-178). This suggests that the flexible terminus may function as a solubility tag, which was supported by the observation by us and others of reduced solubility of *Mt*NusG-constructs lacking parts of this amino-terminal tail (Kalyani et al., 2015).

Structure of MtNusG-CTD

In contrast to the *Mt*NusG-NTD, high-resolution NMR data were easily obtained for the *Mt*NusG-CTD using a construct comprising residues 178-238. Resonance assignments for *Mt*NusG-CTD were possible using standard double- and triple-resonance through-bond and through-space NMR experiments, yielding high-quality spectra (Figures 4(A) and S4). A total of 1016 restraints for the structure calculation of the CTD could be derived from the NMR data (Table 2). The ensemble of 10 structures resulting from the final structure calculation (Figure 4(B)) shows no distance restraint violations larger than 0.1 Å and is devoid of dihedral restraint violations larger than 3°. The structures superimpose

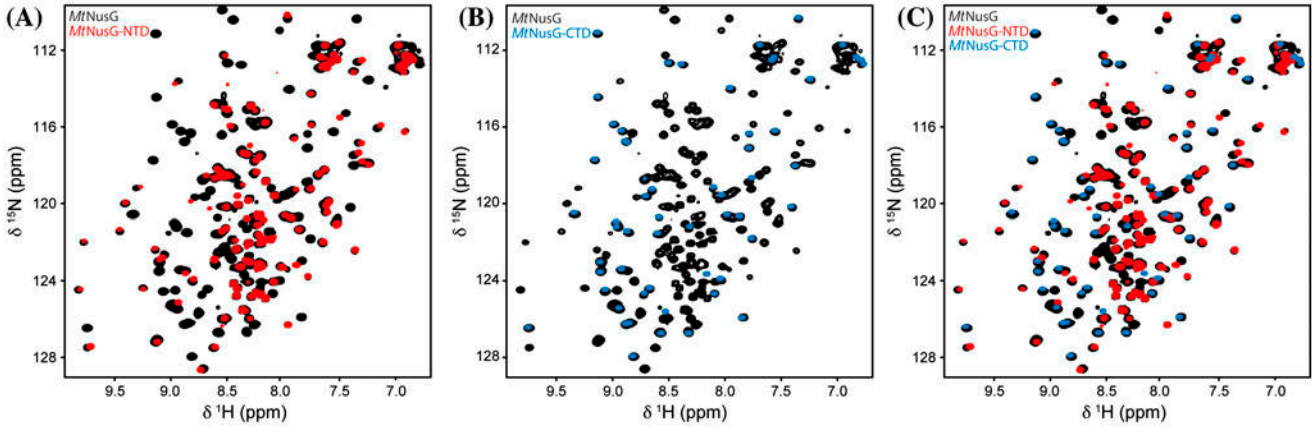


Figure 2. Overlays of the [^{1}H, ^{15}N] HSQC spectrum of *Mt*NusG with the [^{1}H, ^{15}N] HSQC spectra of *Mt*NusG-CTD and *Mt*NusG-NTD. (A) 300 µM [*U*-^{15}N]-*Mt*NusG in 20 mM Na-phosphate, 150 mM NaCl, pH 7.0, black. 700 µM of [*U*-^{15}N]-*Mt*NusG-NTD in 25 mM HEPES, 150 mM NaCl, pH 7.0, red. (B) Spectrum as in (A) black, 1.1 mM [*U*-^{13}C, ^{15}N]-*Mt*NusG-CTD in 25 mM HEPES, 150 mM NaCl, pH 7.5, blue. (C) Overlay of the three spectra. Colors are as in panel A and B. Temperature was 298 K for all measurements.

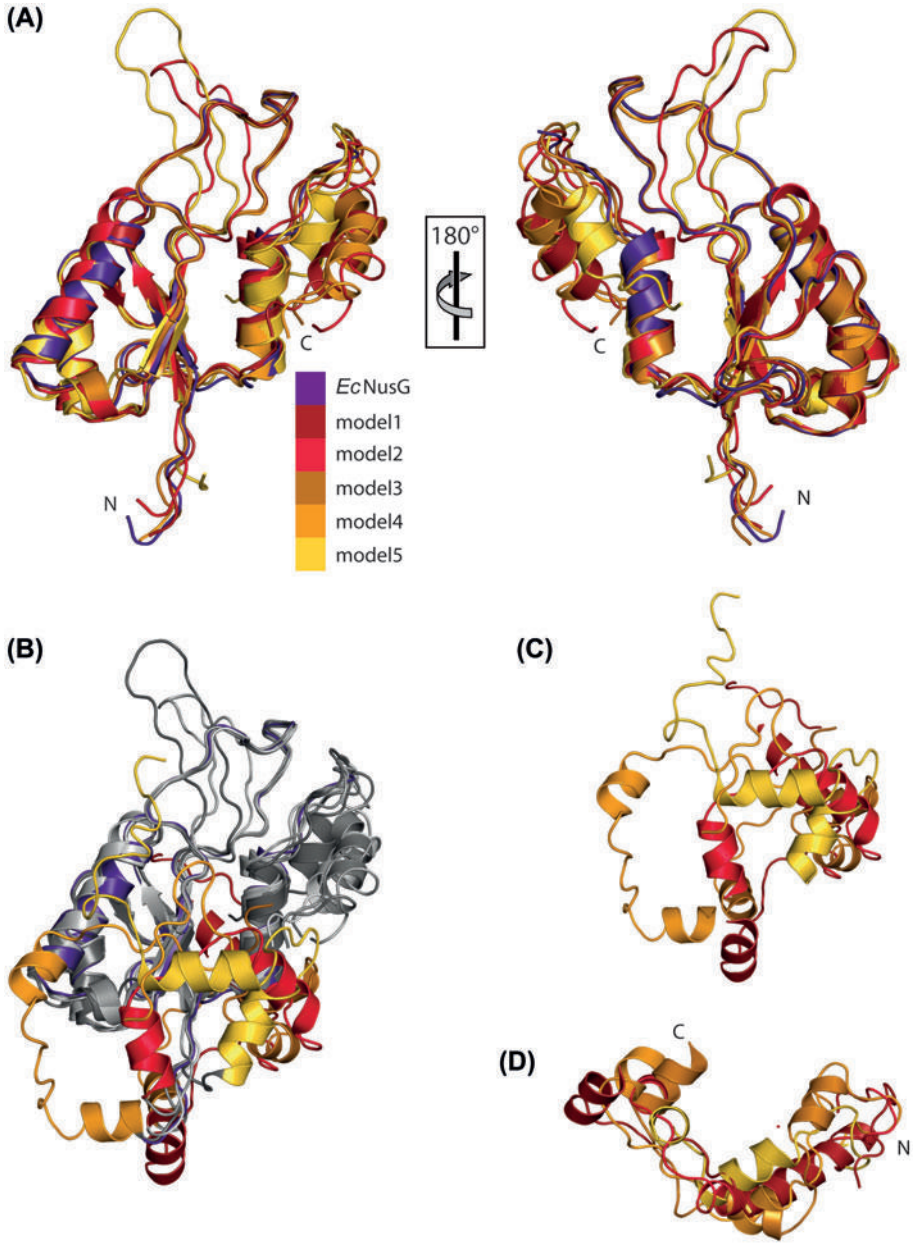


Figure 3. Homology models for *Mt*NusG-NTD calculated via the I-TASSER web server. (A) Cartoon representation of the five *Mt*NusG-NTD models calculated by I-TASSER (Roy et al., 2010) and the structure of *Ec*NusG-NTD (PDB-code: 2K06, Mooney et al., 2009), purple. The initial 40 residues are omitted for clarity. (B) Gray: the five homology models in the same orientation as in (A). The initial 40 amino acids are colored as in (A). These residues show partly helical content, but lack fixed orientation relative to the rest of the NTD and are therefore assumed to be flexible. (C) Isolated initial 40 amino acids. (D) Structural alignment of the first 40 amino acids of the five different models.

well with a coordinate precision of 0.3 Å backbone rmsd (root mean square deviation) and reasonable stereochemical properties.

*Mt*NusG-CTD consists of five β strands (β1: Ser190–Val193; β2: Pro202–Asn209; β3: Lys214–Val219; β4: Thr226–Thr231; and β5: Val235–Ile238, numbering according to wild-type full-length *Mt*NusG) forming an antiparallel barrel-type β-sheet with strand order β5-β1-β2-β3-β4 (Figure 4(C)), in good agreement with the CD data (Table 1, Figure 1(C)). Analysis of the secondary chemical shifts showed that these secondary structure elements are stable in solution (Figure S5). On the basis of an *in silico* model of the *Mt*NusG-CTD, a slightly distorted CTD with altered β-strand orientations

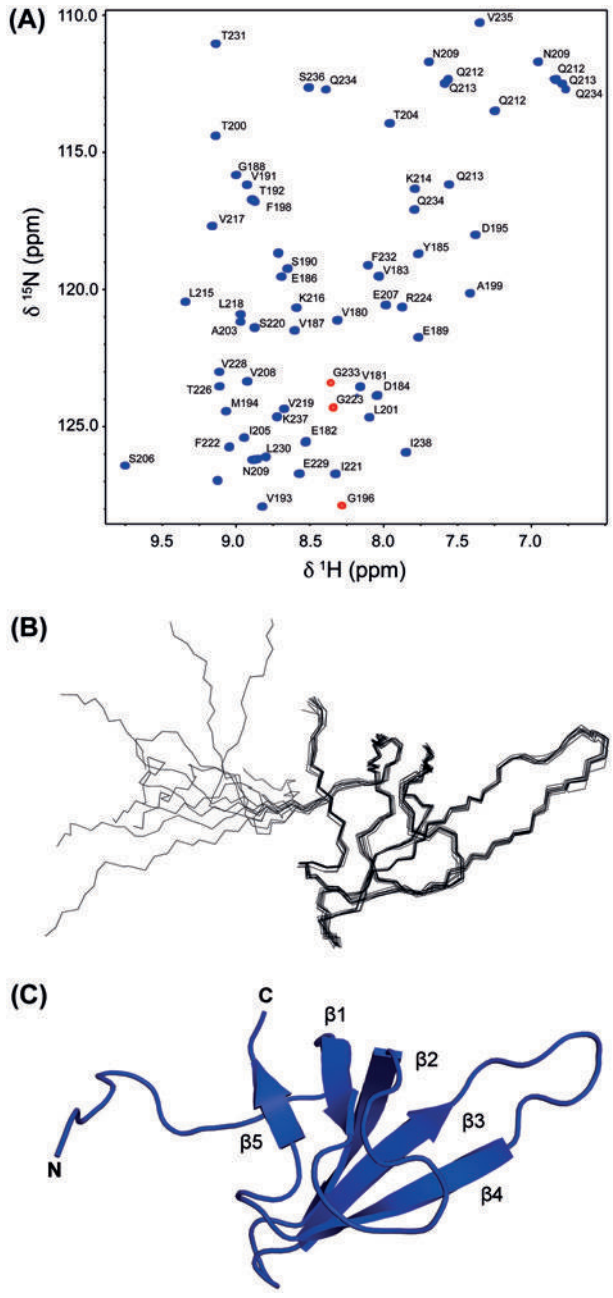


Figure 4. Solution structure of *Mt*NusG-CTD. (A) [^{1}H, ^{15}N] HSQC spectrum of 1.1 mM [*U*-^{13}C, ^{15}N]-*Mt*NusG-CTD in 25 mM HEPES, 150 mM NaCl, and pH 7.5 at 298 K. The sequence-specific resonance assignments obtained from 3D triple resonance experiments are indicated. Folded glycine peaks are shown in red. (B) Structural ensemble of 10 accepted lowest energy structures. The disordered region belongs to the amino-terminal part of the CTD and is part of the flexible linker within the full-length protein. (C) Ribbon representation of the calculated structure exhibiting the lowest energy. The structure was determined with experimental NMR restraints obtained from the *Mt*NusG-CTD construct. The termini as well as the secondary structure elements are indicated.

within the barrel-like CTD structure was proposed (Kalyani et al., 2015). Based on our NOE data and especially on the RDCs, yielding orientational restraints, we can unambiguously show that the *Mt*NusG-CTD exhibits a typical NusG-CTD-fold under our experimental conditions. The *Mt*NusG-CTD solution structure compared to the corresponding domains of *Ec*NusG (PDB code: 2JVV; Mooney et al., 2009), *Aa*NusG (PDB code: 1M1G; Steiner et al., 2002), *Tt*NusG (PDB code: 1NZ9; Reay et al., 2004), and *Tm*NusG (PDB code: 2LQ8; Drögemüller et al., 2013) yielded backbone rmsds of 0.9–1.4 Å, confirming their high degree of structural similarity.

The domains of MtNusG are flexibly linked

An overlay of the two-dimensional (2D) spectra of the full-length protein and *Mt*NusG-CTD (Figure 2(B)) showed that the chemical shifts of the residues belonging to the CTD were virtually identical in the full-length protein and the isolated CTD for all residues within the folded domain. Small deviations could only be observed for the resonances of the amino-terminal residues of the CTD that are part of the linker region between NTD and CTD in the full-length protein. These small chemical shift differences can be attributed to different linker conformations in the wild-type *Mt*NusG compared to the single-domain construct. The 2D [^{1}H, ^{15}N] HSQC-spectrum of *Mt*NusG-NTD also superimposes well with the spectrum of full-length *Mt*NusG (Figure 2(A)). Slight alterations in buffer- and pH-conditions might be the reason for the minor differences in the chemical shifts. To assess the possibility of transient domain interactions, we performed a titration of [*U*-^{15}N]-*Mt*NusG-NTD with increasing amounts of unlabeled *Mt*NusG-CTD up to a twofold excess, however, no chemical shift changes or changes in signal intensities were observed (Figure S6). The absence of spectral changes upon titration of the individual domains is in stark contrast to the effects observed for *Tm*NusG, where domain interactions can clearly be derived from the large chemical shift differences caused by titration of either domain with its complementary domain (Drögemüller et al., 2013).

^{15}N relaxation measurements were performed to further investigate the relative domain motions. The ratios of ^{15}N transversal (R_2) and longitudinal (R_1) relaxation rates are sensitive to the tumbling of proteins. With increasing rotational correlation time, R_1 decreases, while R_2 increases. Therefore, the ratio R_2/R_1 is a sensitive measure of molecular reorientation in solution. Globular proteins are characterized by a single rotational diffusion tensor. This results in a narrow and uniform distribution of R_2/R_1 ratios. In the case of different correlation times for reorientation of the two domains, a non-uniform

Table 2. Solution structure statistics.

Experimentally derived restraints		
Distance restraints		
	NOE	867
	Intraresidual	372
	Sequential	208
	Medium range	59
	Long range	228
	Hydrogen bonds	34
Dihedral restraints		78
Residual dipolar couplings		37
Restraint violation		
Average distance restraint violation (Å)	0.004 ± 0.0002	
Maximum distance restraint violation (Å)	< 0.1	
Average dihedral restraint violation (°)	0.2 ± 0.08	
Maximum dihedral restraint violation (°)	2.6	
Average rdc restraint violation (Hz)	0.09 ± 0.01	
Maximum rdc restraint violation (Hz)	0.32	
Deviation from ideal geometry		
Bond length (Å)	0.00043 ± 0.00002	
Bond angle (°)	0.09 ± 0.002	
Coordinate precision[a,b]		
Backbone heavy atoms (Å)	0.30	
All heavy atoms (Å)	0.76	
Ramachandran plot statistics[c] (%)	89.6/9.2/1.2/0.0	

[a]The precision of the coordinates is defined as the average atomic root mean square difference between the accepted simulated annealing structures and the corresponding mean structure calculated for the given sequence regions.
[b]calculated for residues Asp184-Ile283 (numbering according to full-length *Mt*NusG).
[c]Ramachandran plot statistics are determined by PROCHECK and noted by most favored/additionally allowed/generously allowed/disallowed.

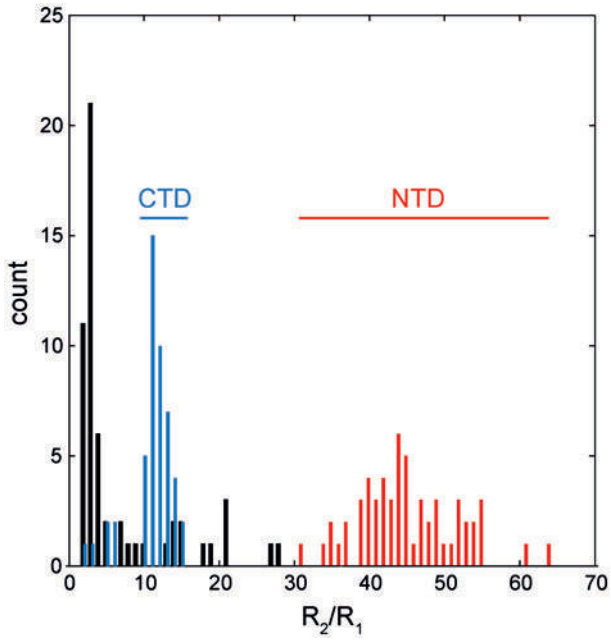


Figure 5. *Mt*NusG domains do not interact. The distribution of the ratio of ^{15}N transverse relaxation rate R_2 and longitudinal relaxation rate R_1 at 16.1 T. Residues from the CTD (blue), NTD (red), and other residues (black) are indicated. The experiments were measured with a sample of 400 µM [*U*-^{2}H, ^{15}N]-*Mt*NusG in 20 mM Na-phosphate, 150 mM NaCl, pH 7.0 at 298 K.

R_2/R_1 distribution will be observed (Burmann et al., 2012; Drögemüller et al., 2013; Horstmann et al., 2006). The distribution of R_2/R_1 ratios of full-length *Mt*NusG is composed of mainly three regions (Figure 5). R_2/R_1 ratios about 1–8 are characteristic for highly flexible regions of the protein, here the amino-terminal extension and the interdomain linker. The region with R_2/R_1 ratios about 10–16 (range of R_1: 1.3–1.4 s^{-1}, range of R_2: 14–20 s^{-1}) corresponds to residues which in large majority belong to the CTD (no CTD residue shows an R_2/R_1 ratio larger than 16). This corresponds to an approximate effective rotational correlation time of 8–12 ns. Residues with R_2/R_1 ratio in the range of 30–60 (range of R_1: 0.7–0.8 s^{-1}, range of R_2: 25–38 s^{-1}) are in the NTD. The significant differences of R_2/R_1 ratios between residues from CTD and NTD demonstrate a different reorientation on the timescale of molecular rotation (low ns) and rule out a defined stable domain interaction in *Mt*NusG. From the relaxation data, approximate effective rotational correlation times of 10 (±2) ns for the CTD and 22 (±5) ns for the NTD are estimated. The presence of large interdomain motions, altering the global shape of the protein during molecular tumbling, prevents a more detailed analysis.

Thus, no tight domain interactions could be detected for *Mt*NusG, in complete analogy to the *E. coli* protein (Burmann et al., 2011).

Conclusion

The data presented here show that the *Mt*NusG-CTD exhibits a typical NusG-CTD-fold as is additionally evidenced by the ability of this domain to interact with *Mt*NusE (Kalyani et al., 2015). This is in contrast to a prediction on the basis of CD-spectroscopy measurements and trypsin proteolytic cleavage (Kalyani et al., 2015). In Kalyani et al. (2015), *Ec*NusG-CTD was prone to trypsin digestion, in contrast to *Mt*NusG-CTD. Therefore, it was proposed that *Mt*NusG-CTD has a different, more compact structure. These results, however, have to be treated with care as the *Ec*NusG-CTD construct used has two potential trypsin cleavage sites at the

amino-terminus, which corresponds to the flexible, unstructured linker in the full-length protein, while their *Mt*NusG-CTD construct has no potential trypsin cleavage site in this region. Therefore, the observed protein cleavage of *Ec*NusG-CTD may be due to a cleavage in this unstructured region.

We also showed that the domains of *Mt*NusG are flexibly linked and do not noticeably interact with each other. Therefore, among the NusG-type proteins studied so far, the thermostable *Tm*NusG is the only one where interactions between NTD and β-barrel-type CTD can easily be detected by [^{1}H, ^{15}N] HSQC (Drögemüller et al., 2013). Functionally, *Tm*NusG seems to be in an autoinhibited state as the CTD:NTD interaction masks the respective binding surfaces to NusE, Rho, and RNAP for both domains, in analogy to the *E. coli* protein RfaH, where, however, the CTD transforms completely to α-helical state to accommodate stable domain interactions (Belogurov et al., 2007; Burmann et al., 2012).

The main function of NusG in *E. coli* is modification of RNAP into a pause-resistant state and the maintenance of operon borders by transcription termination in combination with the Rho termination factor in order to silence horizontally acquired gene (Cardinale et al., 2008). Its homologous protein *Tm*NusG favors forward translocation of RNAP like its *E. coli* counterpart, but slows down transcription elongation like *B. subtilis* NusG (Sevostyanova & Artsimovitch, 2010; Yakhnin & Babitzke, 2010). Initial functional characterization of *B. subtilis* NusG revealed its stimulating effect in hairpin-dependent termination, which resembles *Mt*NusG action (Czyz, Mooney, Iaconi, & Landick, 2014; Kalyani et al., 2015; Yakhnin & Babitzke, 2010). The termination enhancing effect of mycobacterial NusG does neither rely on the electrostatic properties (Figure S3) nor the N-terminal extension nor the enhanced linker (Czyz et al., 2014). In addition, we showed that the structure and domain flexibility of *Mt*NusG is similar to that of *Ec*NusG, leading to the assumption that differences in the binding sites or structural effects on RNAP are the reasons for the functional differences. This, however, has to be elucidated in more detail in future studies. Based on the variable effects of NusG from different organisms, a basic function of NusG proteins as a linker protein for different factors modulating RNAP could be derived (Sevostyanova & Artsimovitch, 2010), whereas the bacterial species-specific functions remain to be functionally and structurally assessed.

Supplementary material

The supplementary material for this paper is available online at http://dx.doi.10.1080/07391102.2015.1031700.

Acknowledgment

We thank Ramona Heißmann and Ulrike Persau for excellent technical support.

Disclosure statement

The authors declare no competing financial interest.

Funding

This work was supported by the Deutsche Forschungsgemeinschaft [grant number Ro 617/17-1].

References

Belogurov, G. A., Vassylyeva, M. N., Svetlov, V., Klyuyev, S., Grishin, N. V., Vassylyev, D. G., & Artsimovitch, I. (2007). Structural basis for converting a general transcription factor into an operon-specific virulence regulator. *Molecular Cell, 26*, 117–129.

Burmann, B. M., Knauer, S. H., Sevostyanova, A., Schweimer, K., Mooney, R. A., Landick, … Rösch, P. (2012). An α helix to β barrel domain switch transforms the transcription factor RfaH into a translation factor. *Cell, 150*, 291–303.

Burmann, B. M., Scheckenhofer, U., Schweimer, K., & Rösch, P. (2011). Domain interactions of the transcription–translation coupling factor *Escherichia coli* NusG are intermolecular and transient. *Biochemical Journal, 435*, 783–789.

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

Cardinale, C. J., Washburn, R. S., Tadigotla, V. R., Brown, L. M., Gottesman, M. E., & Nudler, E. (2008). Termination factor rho and its cofactors NusA and NusG silence foreign DNA in *E. coli*. *Science, 320*, 935–938.

Cornilescu, G., Delaglio, F., & Bax, A. (1999). Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *Journal of Biomolecular NMR, 13*, 289–302.

Czyz, A., Mooney, R. A., Iaconi, A., & Landick, R. (2014). Mycobacterial RNA polymerase requires a U-tract at intrinsic terminators and is aided by NusG at suboptimal terminators. *Mbio, 5*, e00931–14.

DeLano, W. L. (2007). *The PyMOL molecular graphics system*. South San Francisco, CA: DeLano Scientific LLC.

Dolinsky, T. J., Nielsen, J. E., McCammon, J. A., & Baker, N. A. (2004). PDB2PQR: An automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Research, 32*, W665–W667.

Drögemüller, J., Stegmann, C. M., Mandal, A., Steiner, T., Burmann, B. M., Gottesman, M. E., … Schweimer, K. (2013). An autoinhibited state in the structure of thermotoga maritima NusG. *Structure, 21*, 365–375.

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D., & Bairoch, A. (2005). Protein identification and analysis tools on the ExPASy server. In *The proteomics handbook* (pp. 571–607). New York, NY: Humana Press.

Hansen, M. R., Mueller, L., & Pardi, A. (1998). Tunable alignment of macromolecules by filamentous phage yields dipolar coupling interactions. *Nature Structural Biology, 5*, 1065–1074.

Horstmann, M., Ehses, P., Schweimer, K., Steinert, M., Kamphausen, T., Fischer, G., … Faber, C. (2006). Domain motions of the mip protein from *Legionella pneumophila*. *Biochemistry, 45*, 12303–12311.

Johnson, B. A. (2004). Using NMRView to visualize and analyze the NMR spectra of macromolecules. *Methods in Molecular Biology, 278*, 313–352.

Jones, D. T. (1999). Protein secondary structure prediction based on position-specific scoring matrices. *Journal of Molecular Biology, 292*, 195–202.

Kalyani, B. S., Kunamneni, R., Wal, M., Ranjan, A., & Sen, R. (2015). A NusG paralogue from *Mycobacterium tuberculosis*, Rv0639, has evolved to interact with ribosomal protein S10 (Rv0700) but not to function as a transcription elongation-termination factor. *Microbiology, 161*, 67–83.

Kay, L. E., Torchia, D. A., & Bax, A. (1989). Backbone dynamics of proteins as studied by nitrogen-15 inverse detected heteronuclear NMR spectroscopy: Application to staphylococcal nuclease. *Biochemistry, 28*, 8972–8979.

Knowlton, J. R., Bubunenko, M., Andrykovitch, M., Guo, W., Routzahn, K. M., Waugh, D. S., … 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.

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., … Higgins, D. G. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics, 23*, 2947–2948.

Laskowski, R. A., MacArthur, M. W., Moss, D. S., & Thornton, J. M. (1993). PROCHECK: A program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography, 26*, 283–291.

Lescop, E., Schanda, P., & Brutscher, B. (2007). A set of BEST triple-resonance experiments for time-optimized protein resonance assignment. *Journal of Magnetic Resonance, 187*, 163–169.

Liao, D., Lurz, R., Dobrinski, B., & Dennis, P. P. (1996). A NusG-like protein from *Thermotoga maritima* binds to DNA and RNA. *Journal of Bacteriology, 178*, 4089–4098.

Meyer, O., & Schlegel, H. G. (1983). Biology of aerobic carbon monoxide-oxidizing bacteria. *Annual Review of Microbiology, 37*, 277–310.

Mooney, R. A., Schweimer, K., Rösch, P., Gottesman, M. E., & Landick, R. (2009). Two structurally independent domains of *E. coli* NusG create regulatory plasticity via distinct interactions with RNA polymerase and regulators. *Journal of Molecular Biology, 391*, 341–358.

Ottiger, M., Delaglio, F., & Bax, A. (1998). Measurement of *J* and dipolar couplings from simplified two-dimensional NMR spectra. *Journal of Magnetic Resonance, 131*, 373–378.

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

Roberts, J. W., Shankar, S., & Filter, J. J. (2008). RNA polymerase elongation factors. *Annual Review of Microbiology, 62*, 211–233.

Roy, A., Kucukural, A., & Zhang, Y. (2010). I-TASSER: A unified platform for automated protein structure and function prediction. *Nature Protocols, 5*, 725–738.

Sambrook, J., Fritsch, E. F., & Maniatis, T. (1994). *Molecular cloning – A laboratory manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Schwieters, C. D., Kuszewski, J. J., Tjandra, N., & Clore, G. M. (2003). The Xplor-NIH NMR molecular structure determination package. *Journal of Magnetic Resonance, 160*, 66–74.

Sevostyanova, A., & Artsimovitch, I. (2010). Functional analysis of *Thermus thermophilus* transcription factor NusG. *Nucleic Acids Research, 38*, 7432–7445.

Sreerama, N., & Woody, R. W. (2000). Estimation of protein secondary structure from circular dichroism spectra: Comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. *Analytical Biochemistry, 287*, 252–260.

Steiner, T., Kaiser, J. T., Marinkovic, S., Huber, R., & Wahl, M. C. (2002). Crystal structures of transcription factor NusG in light of its nucleic acid- and protein-binding activities. *The EMBO Journal, 21*, 4641–4653.

Werner, F. (2012). A nexus for gene expression – Molecular mechanisms of Spt5 and NusG in the three domains of life. *Journal of Molecular Biology, 417*, 13–27.

Yakhnin, A. V., & Babitzke, P. (2010). Mechanism of NusG-stimulated pausing, hairpin-dependent pause site selection and intrinsic termination at overlapping pause and termination sites in the *Bacillus subtilis trp* leader. *Molecular Microbiology, 76*, 690–705.

Supplementary Information

The two domains of *Mycobacterium tuberculosis* NusG protein are dynamically independent

Martin Strauß[a,*], Kristian Schweimer[a], Björn M. Burmann[a,b], Anne Richter[a], Stephanie Güttler[c], Birgitta M. Wöhrl[a] and Paul Rösch[a]

[a]Lehrstuhl Biopolymere und Forschungszentrum für Bio-Makromoleküle, Universität Bayreuth, Universitätsstraße 30, 95447 Bayreuth, Germany

[b]Biozentrum, University of Basel, Klingelbergstr. 70, 4056 Basel, Switzerland;

[c]Lehrstuhl Tierökologie II, Universität Bayreuth, Universitätsstraße 30, 95447 Bayreuth, Germany

Contents

Figure S1:

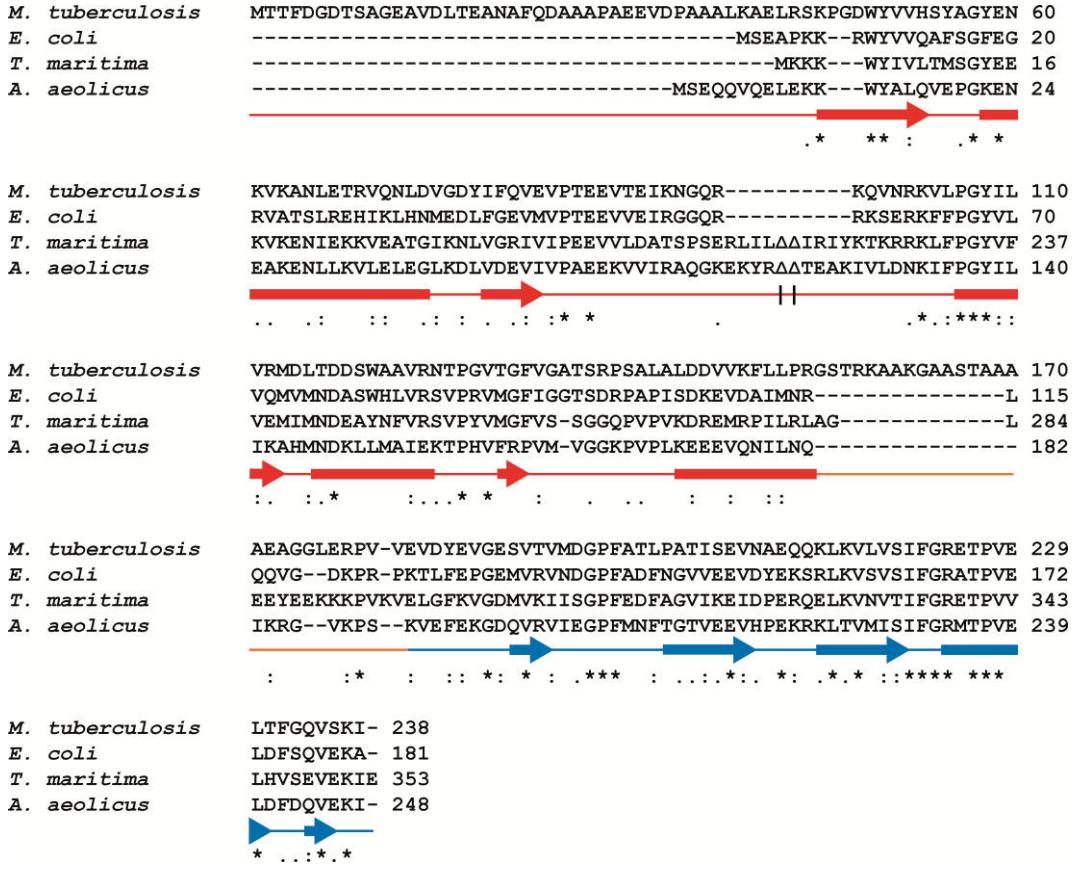


Figure S1. Alignment of sequences of NusG proteins from *M. tuberculosis*, *E. coli* , *T. maritima,* and *A. aeolicus*. For *T. maritima* and *A. aeolicus* NusG the residues corresponding to the additional domain marked ΔΔ were not considered (*T. maritima*: Δ58-120; *A. aeolicus*: Δ66-125). The secondary structure elements on the basis of the *Tm*NusG crystal structure are indicated below the sequences: arrows, β-strands; bars, α-helices. NTD, red; CTD, blue; linker, orange. The alignment was performed with ClustalW2 (Larkin *et al.*, 2007).

Figure S2:

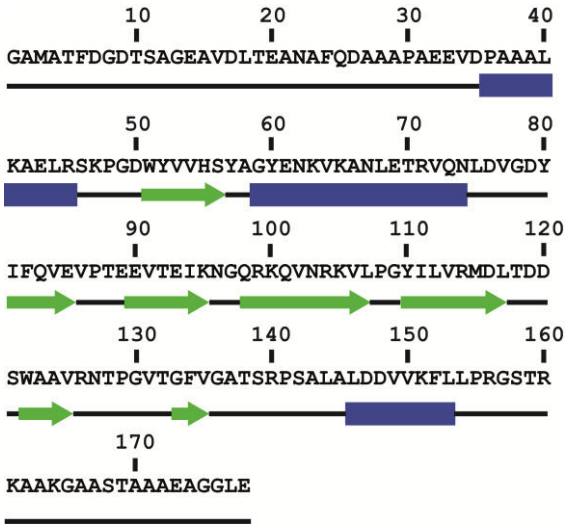


Figure S2. Secondary structure prediction for *Mt*NusG-NTD with PSIPRED (Jones, 1999). α-helices, blue; β-strands, green; unstructured regions, black.

Figure S3:

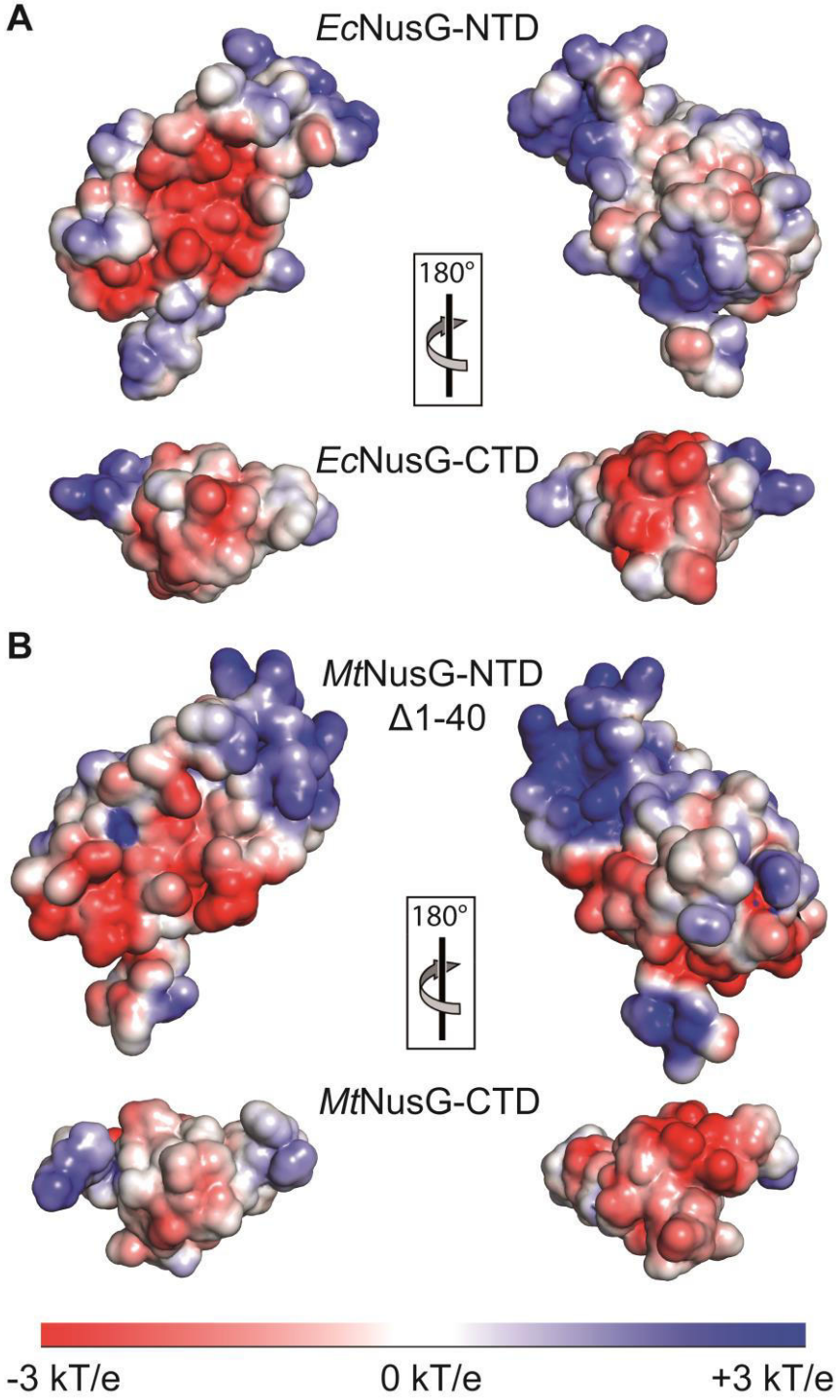


Figure S3. Comparison of the electrostatic potentials on the solvent accessible surfaces of *Mt*NusG and *Ec*NusG. Electrostatic surface potential from -3 kT/e, red, to +3 kT/e, blue. **A)** *E. coli* protein, pdb-codes 2K06 (*Ec*NusG-NTD) and 2JVV (*Ec*NusG-CTD; Mooney, Schweimer, Rösch, Gottesman, & Landick, 2009). **B)** *Mt*NusG-CTD, pdb-code 2MI6, our data, and *Mt*NusG-NTD, first I-TASSER model omitting 40 amino-terminal amino acids.

Figure S4:

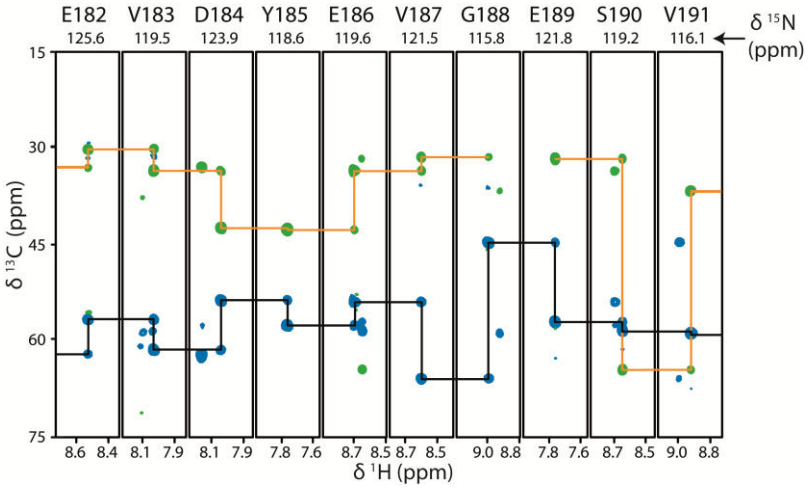


Figure S4. Sequence specific resonance assignments. Representative backbone assignment strips from a 3D BEST-HNCACB experiment (Lescop, Schanda, & Brutscher, 2007).

Figure S5:

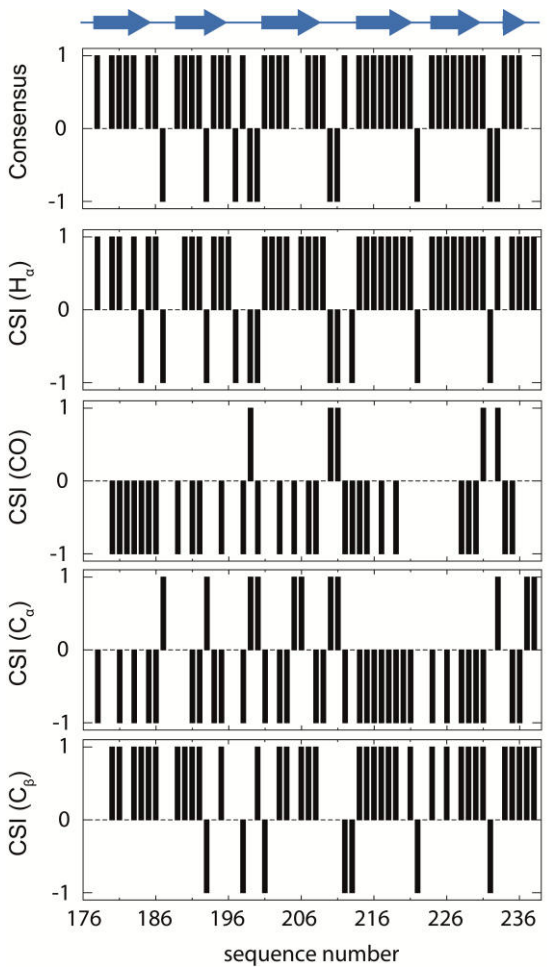


Figure S5. Chemical Shift Index (CSI). CSI for H_α, CO, C_α, and C_β as well as the resulting consensus CSI for the *Mt*NusG-CTD. The secondary structure elements of the *Mt*NusG-CTD are indicated on top.

Figure S6:

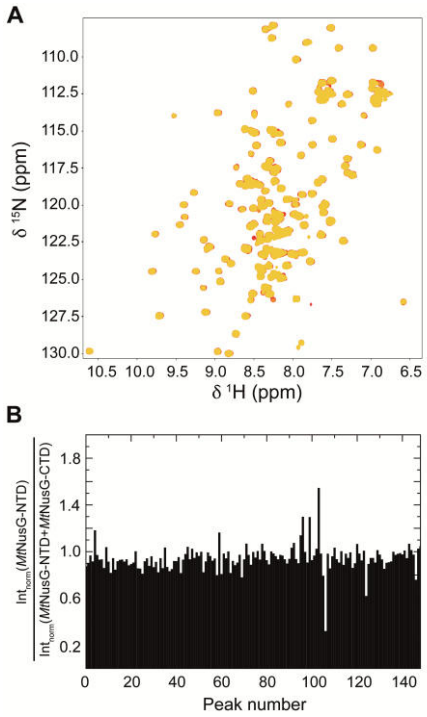


Figure S6: The free *Mt*NusG domains do not interact. A) [^{1}H, ^{15}N] HSQC spectrum of 150 µM [*U*-^{15}N]-*Mt*NusG-NTD in 25 mM HEPES, 150 mM NaCl, pH 7.5, red. Spectra after addition of one or two equivalents *Mt*NusG-CTD, orange and yellow, respectively. Temperature: 298 K. **B)** Relative intensities of the *Mt*NusG-NTD resonances in the presence of two *Mt*NusG-CTD equivalents versus unassigned resonances of free *Mt*NusG-NTD. The intensities were normalized by the differences in 90° pulse-length, and sample dilution.

References

Jones, D. T. (1999). Protein secondary structure prediction based on position-specific scoring matrices. *Journal of Molecular Biology, 292*, 195-202.

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson T. J., Higgins, D. G. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics, 23*, 2947-2948.

Lescop, E., Schanda, P., & Brutscher, B. (2007). A set of BEST triple-resonance experiments for time-optimized protein resonance assignment. *Journal of Magnetic Resonance, 187*, 163-169.

Mooney, R. A., Schweimer, K., Rösch, P., Gottesman, M. E., & Landick, R. (2009). Two structurally independent domains of *E. coli* NusG create regulatory plasticity *via* distinct interactions with RNA polymerase and regulators. *Journal of Molecular Biology, 391*, 341-358.

7.2 Einzelarbeit B

Martin Strauß, Christal Vitiello, Kristian Schweimer, Max E. Gottesman, Paul Rösch, Stefan H. Knauer (2016): **Transcription is regulated by NusA:NusG interaction**, *Nucleic Acids Research,* Mai 2016.

Nucleic Acids Research Advance Access published May 12, 2016

Nucleic Acids Research, 2016 **1**
doi: 10.1093/nar/gkw423

Transcription is regulated by NusA:NusG interaction

Martin Strauß[1], Christal Vitiello[2], Kristian Schweimer[1], Max Gottesman[2,3], Paul Rösch[1] and Stefan H. Knauer[1,*]

[1]Lehrstuhl Biopolymere und Forschungszentrum für Bio-Makromoleküle, Universität Bayreuth, 95447 Bayreuth, Germany, [2]Department of Microbiology and Immunology, Columbia University Medical Center, New York, NY 10032, USA and [3]Department of Biochemistry and Molecular Biophysics, Columbia University Medical Center, New York, NY 10032, USA

Received August 26, 2015; Revised May 04, 2016; Accepted May 05, 2016

ABSTRACT

NusA and NusG are major regulators of bacterial transcription elongation, which act either in concert or antagonistically. Both bind to RNA polymerase (RNAP), regulating pausing as well as intrinsic and Rho-dependent termination. Here, we demonstrate by nuclear magnetic resonance spectroscopy that the *Escherichia coli* NusG amino-terminal domain forms a complex with the acidic repeat domain 2 (AR2) of NusA. The interaction surface of either transcription factor overlaps with the respective binding site for RNAP. We show that NusA-AR2 is able to remove NusG from RNAP. Our *in vivo* and *in vitro* results suggest that interaction between NusA and NusG could play various regulatory roles during transcription, including recruitment of NusG to RNAP, resynchronization of transcription:translation coupling, and modulation of termination efficiency.

INTRODUCTION

Transcription, the first step in gene expression, is highly regulated by a multitude of transcription factors. The core transcription machinery is RNA polymerase (RNAP), which consists of five subunits in bacteria (2 x α, β, β', and ω) (1). RNAP initiates RNA synthesis at a promoter (initiation), extends the nascent RNA (elongation), and releases the RNA at a terminator (termination) (2). Among *Escherichia coli (E. coli)* transcription factors, N-Utilization Substances (Nus) A and G are the only ones known to affect both the speed of RNA chain elongation and termination (3). NusG (Spt5 in archaea and eukaryotes), the only universally conserved transcription factor, is composed of an N-terminal domain (NTD) flexibly connected to a C-terminal domain (CTD) (Supplementary Figure S1A) (4,5). NusG-NTD interacts with the β' clamp helices (β'CH) and the β gate loop (βGL) of RNAP to increase RNAP processivity (4,6,7). NusG-CTD is target of at least two cellular partners, termination factor Rho and antitermination factor NusE, which is identical to ribosomal protein S10 (8,9). NusA comprises an NTD that binds to the β flap tip helix of RNAP at the RNA exit channel, three RNA binding motifs (S1, KH1, KH2) that together form the SKK domain, and, in *E. coli* and other γ-proteobacteria, two additional C-terminal acidic repeat domains (AR1, AR2; Supplementary Figure S1B) (10–14). NusA-AR1 interacts with N protein of phage λ, but is not essential for N-mediated suppression of transcription termination (antitermination) (15–17). NusA-AR2 can bind either to the CTD of the RNAP α-subunit (αCTD) or to the NusA-SKK. NusA-AR2 attached to NusA-SKK autoinhibits NusA activity by preventing RNA binding (15,18).

NusA and NusG differentially alter the properties of the transcription elongation complex (TEC) *via* direct and independent interactions (3). NusG increases TEC processivity whereas NusA slows RNAP by either increasing pause times or by introducing new pause sites (19). The two factors have context-dependent effects on termination and act either in concert or as antagonists. On the one hand, NusG and NusA are proposed to support Rho cooperatively to suppress the toxic functions of foreign genes. On the other hand, both are part of antitermination complexes on ribosomal RNA and phage λ nascent transcripts (20–23). Furthermore, NusA can enhance or decrease both Rho-dependent and intrinsic termination efficiency, depending on the specific terminator (reviewed in (24,25)). NusG, in contrast, enhances termination exclusively at Rho-dependent sites (26,27). Importantly, NusG serves as the physical linker between the RNAP and the ribosome by binding RNAP *via* NusG-NTD and S10 *via* NusG-CTD, thus coupling transcription and translation (8).

NusA and NusG bind to different sites on RNAP (6,7,10). Although these sites are in close proximity, a direct connection between the two factors has not been reported thus far. With nuclear magnetic resonance (NMR) spectroscopy we here demonstrate that NusA and NusG do

*To whom correspondence should be addressed. Tel: +49 921 553868; Fax: +49 921 16490459; Email: stefan.knauer@uni-bayreuth.de
Present address: Martin Strauß, Department of Microbiology and Immunology, Columbia University Medical Center, New York, NY 10032, USA.

specifically mutually interact and, supported by *in vivo* and *in vitro* data, we propose that this interaction may have key regulatory roles in diverse steps of transcription.

MATERIALS AND METHODS

Cloning

nusA was cloned into the pTKK19 expression vector (28) *via* NdeI and BamHI restriction sites resulting in the recombinant plasmid pTKK19_nusA(1-495). The recombinant protein carried a deca-histidine tag followed by a PreScission cleavage site at its N-terminus.

Gene expression and protein purification

Gene expression and protein purification procedures for NusG, NusG-NTD, NusG-CTD, NusA-NTD, NusA-SKK, NusA-AR1 and αCTD were described earlier (8,16,29–32). NusA-AR2 (NusA(424–495)) was produced as fusion protein with His_{10} tag followed by PreScission protease cleavage site at its N-terminus. Its gene expression and protein purification were according to the protocol for NusA(339-495) (33). His_{10}-NusA-AR2 was purified like NusA-AR2 omitting the tag-removal step.

nusA was expressed in *E. coli* BL21 (DE3) harboring pTKK19_nusA(1-495). Cells were grown in lysogeny broth (LB) medium containing 30 μg/ml kanamycin at 37°C. At an optical density at 600 nm (OD_{600}) of ∼0.7 expression was induced by 1 mM isopropyl-thiogalactoside (IPTG). Cells were harvested after 4 h (9,000 x *g*, 15 min, 4°C), resuspended in buffer A (20 mM Tris(hydroxymethyl)aminomethane (Tris)/HCl, pH 7.9, 500 mM NaCl, 5 mM imidazole, 1 mM β-mercaptoethanol) and disrupted by a microfluidizer (Microfluidics, Newton, MA, USA). After centrifugation (12,000 x *g*, 30 min, 4°C) the crude extract was applied to a 5 ml HisTrap HP column (GE Healthcare, Munich, Germany) and eluted using a step gradient from 5 mM to 1 M imidazole in buffer A. Fractions containing the His_{10}-NusA fusion protein were combined and the protein was cleaved by PreScission protease during dialysis against buffer B (20 mM Tris/HCl, pH 8, 1 mM β-mercaptoethanol) at 4°C overnight. The protein solution was applied to a 5 ml GSTrap FF column (GE Healthcare, Munich, Germany) and the flow-through subsequently to a 5 ml QXL column (GE Healthcare, Munich, Germany). NusA was eluted using a step gradient from 0 to 1 M NaCl in buffer B. Fractions containing pure NusA were combined, dialyzed against 5 l 20 mM Tris/HCl, pH 7.5, 50 mM NaCl, 1 mM dithiothreitol (DTT), concentrated using ultrafiltration units (Viva Science, molecular weight cut-off (MWCO): 10 kDa), shock frozen in liquid nitrogen, and stored at −80°C.

Production and purification of RNAP for NMR experiments was based on Ref. (34). In brief, *E. coli* BL21(DE3) (Novagen, Madison, WI, USA) harboring a plasmid containing *rpoA*, *rpoB*, *rpoC* and *rpoZ* in one operon under control of T7 promoter was used for gene expression. The β' subunit was produced as fusion protein carrying a His_6 tag at its C-terminus. 2 l of LB medium in a 5 l flask supplemented with 100 μg/ml ampicillin were inoculated with an overnight culture to an OD_{600} of 0.02 and incubated at 37°C and 150 rpm. Having reached an OD_{600} of 0.2 the temperature was decreased to 20°C. After 2 h, IPTG was added to a final concentration of 0.5 mM for induction, and the culture was incubated overnight. Cells were then harvested by centrifugation for 15 min at 4°C and 9,000 x *g*, resuspended in buffer C (50 mM Tris/HCl, pH 6.9, 0.5 M NaCl, 5 % (v/v) glycerol, 1 mM DTT) and lysed with a microfluidizer (Microfluidics, Newton, MA, USA). RNAP was purified by nickel affinity chromatography with 2×5 ml Ni^{2+}-nitrilotriacetic acid (NTA) Superflow columns (QIAGEN) and eluted by a constant gradient from 0 to 1 M imidazole in buffer C. Peak fractions containing RNAP were dialyzed against buffer D (50 mM Tris/HCl, pH 6.9, 0.5 mM ethylenediaminetetraacetic acid (EDTA), 5 % (v/v) glycerol, 1 mM DTT) overnight at 4°C and then applied to a 5 ml Heparin HP column (GE Healthcare), followed by elution with a constant NaCl gradient from 0 to 1 M NaCl in buffer D. Fractions containing RNAP were pooled, dialyzed against buffer B overnight at 4°C, applied to a 25 ml Q-Sepharose FF column, and eluted with a constant gradient from 0 to 1 M NaCl in buffer B. RNAP containing fractions were concentrated using ultrafiltration units (Viva Science, MWCO: 10 kDa), shock frozen in liquid nitrogen, and stored at −80°C.

Production and purification of RNAP for *in vitro* transcription assays was carried out as described (35), with minor modifications.

Isotopic labeling of proteins

Proteins were uniformly labeled with ^{15}N or ^{15}N,^{13}C by growing *E. coli* in M9 minimal medium (36,37) supplemented with $(^{15}NH_4)_2SO_4$ (Campro Scientific, Berlin, Germany) or $(^{15}NH_4)_2SO_4$ and ^{13}C-D-glucose (Cambridge Isotope laboratories, Inc., Andover, MA, USA) as the only nitrogen and carbon source, respectively. Expression and purification procedures were identical to those used for proteins produced in LB medium.

Pull-down assay

The pull-down assay was performed with a 1 ml HisTrap HP column (GE Healthcare) equilibrated with buffer E (10 mM potassium phosphate, pH 6.9, 50 mM NaCl). The proteins were dialyzed against buffer E overnight at 4°C before application to the column. The application volume was always 1 ml. After extensive washing with buffer E, elution was carried out with 100 or 400 mM imidazole in buffer E, respectively, and resulted in a mixture of His_{10}-NusA-AR2 and NusG-NTD (400 μM and 200 μM, molar ratio: 2:1) as assayed, with His_{10}-NusA-AR2 (200 μM) alone and NusG-NTD (400 μM) alone as controls.

In vitro transcription assay

RNA and DNA oligonucleotides were commercially synthesized by Integrated DNA Technologies with sequences derived from the T7A1 promoter sequence. Assembly of the TEC and the *in vitro* transcription were carried out as described (38). Briefly, a 65mer template DNA strand was hybridized to an 11mer RNA labeled with ^{32}P at the 5′

end. 50 pmol RNAP in transcription buffer (TB, 20 mM Tris/HCl, pH 7.9, 5 mM $MgCl_2$, 40 mM KCl, 2 mM β-mercaptoethanol) were mixed with equimolar concentrations of the DNA:RNA hybrid, followed by addition of the nontemplate DNA strand. The assembled TECs were purified by affinity chromatography using Ni^{2+}-nitrilotriacetic acid agarose (QIAGEN) and subsequent membrane filtration with Ultrafree© 0.65 μm PVDF centrifugal filters (Millipore). 2.5 μM TEC were incubated with 50 μM of one or both transcription factors or TB for 10 min at 25°C. When two proteins were tested they were added simultaneously, as previous tests indicated that the order of addition had no effect. Transcription was initiated by addition of 1 mM NTPs and stopped after 60 seconds by addition of an equal volume of 2x loading buffer (10M urea, 50 mM EDTA, pH 7.9, 0.05 % (w/v) bromophenol blue, and 0.05 % (w/v) xylene cyanol). RNA products were resolved on a 23 % denaturing polyacrylamide gel containing 7 M urea. Gels were exposed to phosphor screens and scanned by Typhoon Phosphorimager (GE Healthcare Life Sciences).

NMR experiments

All NMR spectra were recorded at 298 K on Bruker Avance 700 MHz and Avance 800 MHz spectrometers with cryogenically cooled triple-resonance probes equipped with pulsed field-gradient capabilities. Processing of NMR data was carried out using in-house routines and visualized by NMRView (39). For all NMR experiments the proteins were in 10 mM potassium phosphate buffer, pH 6.4, 50 mM NaCl, at 298 K. The initial sample volume was 550 μl if not stated otherwise. Backbone assignments of NusG-NTD, NusG-CTD, NusA-AR2, and αCTD were taken from previous studies (4,18,32).

To evaluate [^{1}H,^{15}N]-heteronuclear single quantum coherence (HSQC) titration experiments we calculated the normalized chemical shift changes ($\Delta\delta_{norm}$) according to Equation (1).

$$\Delta\delta_{norm} = \sqrt{\Delta\delta(^1\mathrm{H})^2 + \left[0.1 \cdot \Delta\delta\,(^{15}\mathrm{N})\right]^2} \qquad (1)$$

Dissociation constants (K_D) were calculated from [^{1}H,^{15}N]-HSQC titrations by analyzing the chemical shift changes and fitting a two-state model as in Equation (2) to the chemical shift change of amide protons showing fast exchange on the chemical shift timescale.

$$\Delta\nu = \Delta\nu_{End} \cdot \frac{[P]_0 \cdot r + [P]_0 + K_D - \sqrt{(K_D + [P]_0 + [P]_0 \cdot r)^2 - 4 \cdot ([P]_0)^2 \cdot r}}{2 \cdot [P]_0} \qquad (2)$$

where $\Delta\nu$ is the normalized resonance frequency difference in Hz, $\Delta\nu_{End}$ the normalized resonance frequency difference between free and fully bound protein in Hz, K_D the dissociation constant, r the protein:labeled protein ratio and $[P]_0$ the total concentration of ^{15}N-labeled protein. K_D and $\Delta\nu_{End}$ were used as fitting parameters. The reduction of $[P]_0$ due to dilution was accounted for during fitting.

For the displacement experiment of ^{15}N-NusA-AR2 from αCTD by NusG-NTD separate samples were prepared for ^{15}N-NusA-AR2 (100 μM, 500 μl) and ^{15}N-NusA-AR2: αCTD (100 μM each, 500 μl). NusG-NTD was then added to the latter sample from a 287 μM stock solution.

For the quantitative analysis of signal intensities in the displacement experiments signal intensities were normalized by the number of scans, the concentration, and the length of the 90° proton pulse.

Docking

The complex of NusA-AR2 and NusG-NTD was modeled with the HADDOCK webserver (40) using data from the [^{1}H,^{15}N]-HSQC titrations as restraints (active residues in NusG-NTD: 4, 13, 15, 18, 44, 50, 51, 52, 58, 59, 61, 95, 104, 105, 106, 107, 109, 114, 117 and 118; active residues in NusA-AR2: 463, 474, 483, 487, 489, 490, 491 and 493). Passive residues were determined automatically. The NMR ensembles of NusA-AR2 (Protein Data Bank (PDB) ID: 1WCN) and NusG-NTD (PDB ID: 2K06) were used as input.

Strains and β-galactosidase assays

Strains were derivatives of MDS42, which lacks prophages and insertion elements (41) containing λ fusions (42). *lacZ* is expressed from the fusions *λcI857 – pR – cro(ΔRBS) – nutR – tR1 – cII::lacZ* or *λcI857 – pR – cro$_{27}$ – nutR – tR1 – cII::lacZ*. The TAAGGAGGTTGT to TaccctccTTGT substitution in the *cro* ribosome binding site (RBS), blocks *cro* translation in the former strain (*cro(ΔRBS)*). *cro$_{27}$* has an RBS, but Cro_{27} is non functional due to an amino acid exchange (R27Q). The creation of strains 10323 and 10881 was described previously (18). The *rhoE134K* mutation was introduced by phage P1 transduction, resulting in a Rho variant which is non functional at *λtR1*. The strains carrying *nusA* variants were constructed by recombineering.

Cells were assayed for β-galactosidase activity (Miller units) after overnight growth at 37°C. β-galactosidase activity of cells with defective Rho (*rho15*) was set to 100 % (strains 11633 and 11634), since termination was completely abolished at *λtR1*. Assays were performed independently four times and resulting activities were averaged.

Programmes

Graphical representations of protein structures were created with PyMOL (43). Sequence alignments were done with Clustal omega (44).

RESULTS AND DISCUSSION

NusG interacts specifically with NusA

First, we probed a possible NusA:NusG interaction by NMR spectroscopy with full-length proteins. Addition of NusA to ^{15}N-labeled NusG to equimolar concentration resulted in a strong decrease of NusG-NTD signals in the [^{1}H,^{15}N]-HSQC spectrum, whereas NusG-CTD signals were weakened only marginally (Figure 1A). The high transversal relaxation rate of the 54.9 kDa NusA strongly affects magnetization transfer efficiency upon binding, which leads to line broadening and ultimately to a decrease of signal intensity. Thus, the observed loss of NusG-NTD signals suggests direct NusA:NusG-NTD interaction. Specific NusA:NusG-NTD complex formation was confirmed

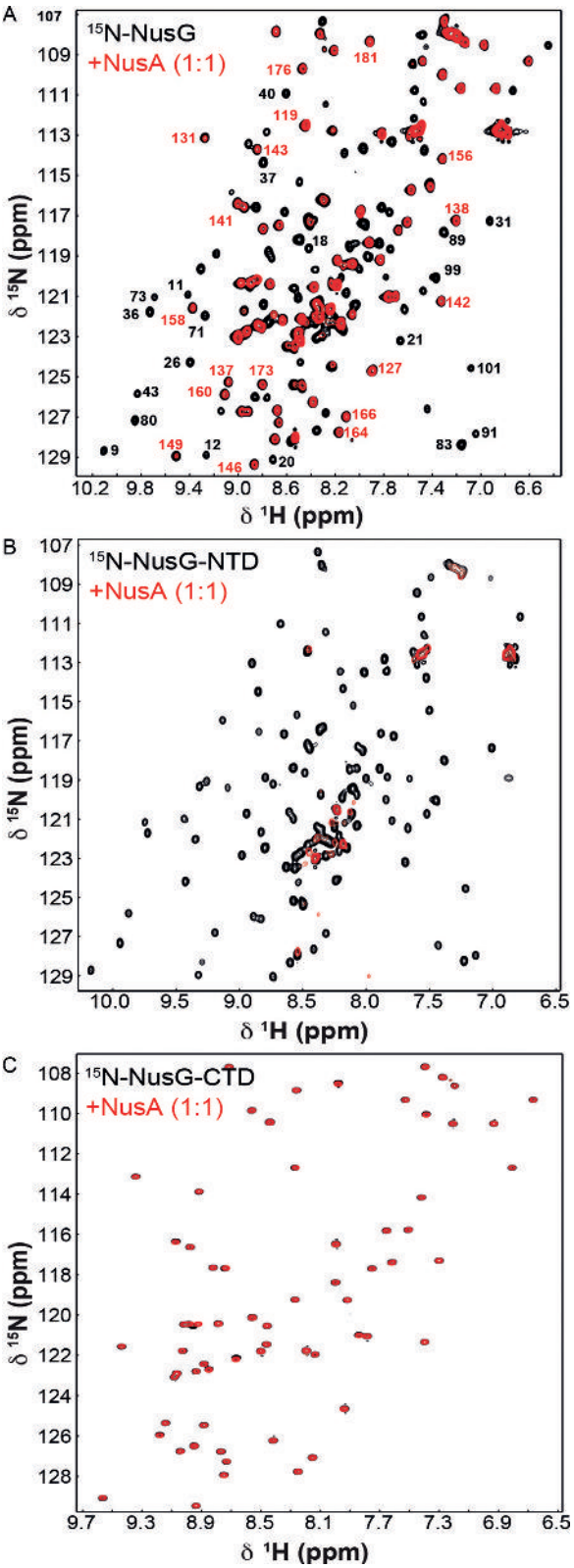


Figure 1. NusG-NTD interacts with NusA. [^{1}H,^{15}N]-HSQC spectra of (**A**) ^{15}N-NusG, (**B**) ^{15}N-NusG-NTD and (**C**) ^{15}N-NusG-CTD before, black, and after, red, addition of NusA in equimolar concentration. Numbers in (**A**) represent the corresponding amino acid number of NusG with NusG-NTD signals in red and NusG-CTD signals in black. The concentration of the NusG construct was 50 μM in all experiments.

by addition of NusA to either isolated ^{15}N-NusG-NTD or ^{15}N-NusG-CTD as signal loss was only observed for ^{15}N-NusG-NTD (Figure 1B and C). [^{1}H,^{15}N]-HSQC spectra of a 2-fold molar excess of ^{15}N-NusG-NTD in the presence of isolated NusA-NTD, NusA-SKK, NusA-AR1, or NusA-AR2 clearly showed that, of these, only NusA-AR2 interacted directly with NusG-NTD (Figure 2).

To corroborate the results of the NMR experiments, a pull-down assay was performed with decahistidine-tagged NusA-AR2 (His_{10}-NusA-AR2) and untagged NusG-NTD (Supplementary Figure S2). NusG-NTD eluted from the nickel column together with His_{10}-NusA-AR2, thus confirming the direct interaction between the two domains.

The NusG-NTD:NusA-AR2 complex

From the chemical shift perturbations of [^{1}H,^{15}N]-HSQC NMR titrations the K_D-value of the NusG-NTD:NusA-AR2 interaction can be estimated to be 24 μM (Figure 3A, Supplementary Figure S3). By mapping the normalized chemical shift changes on the three-dimensional structures of NusG-NTD and NusA-AR2 the interaction interface can be located in the C-terminal part of helix α5 of NusA-AR2, markedly involving W490 and F491 (Figure 3B and C). Although NusA-AR1 and NusA-AR2 share high sequence identity (31.5%) and have virtually identical three-dimensional structures with a root mean square deviation of main chain atoms of 1.2 Å (Supplementary Figure S4), NusG-NTD recognizes NusA-AR2 exclusively (Figure 2). This selectivity can probably be attributed to the presence of W490 and F491 in NusA-AR2, since Leu and Ala are found at corresponding positions in NusA-AR1 (Supplementary Figure S4). Different binding specificities for NusA-AR1 and NusA-AR2 to λ protein N and the αCTD of RNAP have been noted earlier (13,16,18). NusG-NTD signals from residues in the C-terminal helix α3' (aa 104–117) and in the elongated loop region between α1' and β1' (aa 48–68) are mainly affected by the NusG-NTD:NusA-AR2 interaction (Figure 3B and C). Based on these binding surfaces a docking model without conformational rearrangements of the complex was generated (Figure 3C). Remarkably, the NusA-AR2 binding site on NusG-NTD is also involved in the NusG-NTD:RNAP β'CH interaction (Supplementary Figure S5) (7,45). Furthermore, NusA-AR2 residues responsible for NusG-NTD binding are necessary for the interaction with αCTD (Supplementary Figure S6) (18). Thus, formation of the NusG-NTD:RNAP and NusG-NTD:NusA-AR2 complex is mutually exclusive, as is formation of the NusG-NTD:NusA-AR2 and NusA-AR2:RNAP complex.

NusG-NTD:RNAP versus NusG-NTD:NusA-AR2 versus NusA-AR2:RNAP

We asked if NusG-NTD:NusA-AR2 interaction can take place in the presence of RNAP. Therefore a [^{1}H,^{15}N]-HSQC displacement experiment with isolated αCTD was performed (Figure 4A,B). Addition of NusA-AR2 to ^{15}N-αCTD to equimolar concentration induced the chemical shift perturbations of ^{15}N-αCTD signals observed earlier for this interaction (18). Stepwise addition of NusG-NTD

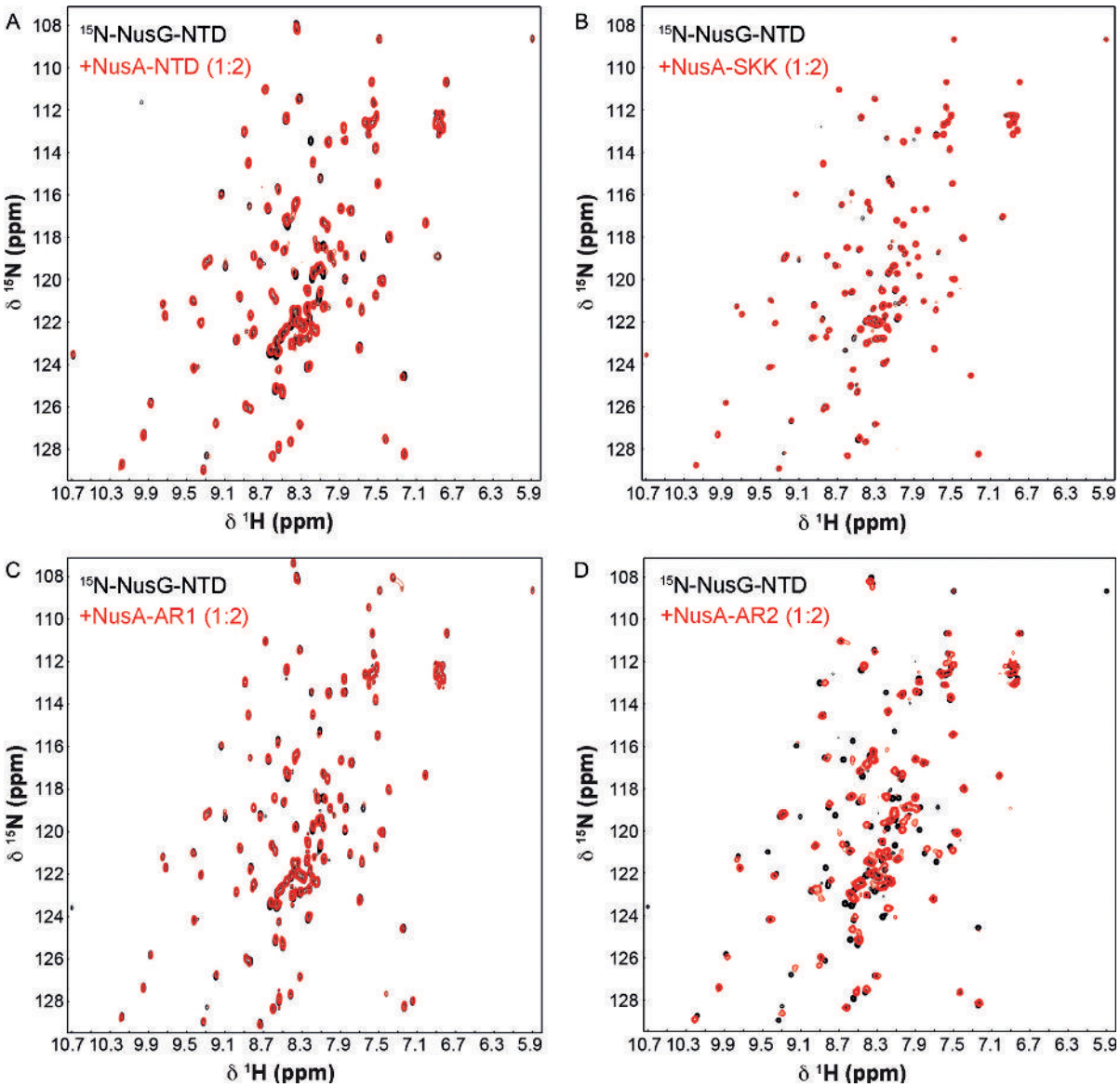


Figure 2. NusG-NTD interacts with NusA-AR2. [^{1}H,^{15}N]-HSQC spectra of ^{15}N-NusG-NTD in the absence, black, or presence, red, of (**A**) NusA-NTD, (**B**) NusA-SKK, (**C**) NusA-AR1, and (**D**) NusA-AR2. ^{15}N-NusG-NTD was present at 100 μM in all experiments, and NusA domains were added in a twofold molar excess.

to a final three-fold molar excess of NusG-NTD partially reversed these shifts, indicating that NusG-NTD can displace NusA-AR2 from the αCTD. Although the displacement was incomplete due to the lower K_D of NusA-AR2:αCTD (K_D < 5 μM) (18) versus NusA-AR2:NusG-NTD (K_D: 24 μM), it confirms that the RNAP/NusA-AR2 binding sites on NusG-NTD overlap as do the αCTD/NusG-NTD interaction interfaces on NusA-AR2. This finding was counter-checked by the displacement of ^{15}N-NusA-AR2 from αCTD by NusG-NTD (Supplementary Figure S7). Addition of αCTD to ^{15}N-NusA-AR2 to equimolar concentration resulted in chemical shift changes of signals from ^{15}N-NusA-AR2 typical for ^{15}N-NusA-AR2:αCTD complex formation (18). Subsequent addition of NusG-NTD caused the ^{15}N-NusA-AR2 signals to shift towards the resonance positions of the ^{15}N-NusA-AR2:NusG-NTD complex. Again, the displacement was incomplete owing to the difference in the affinities of NusA-AR2 to αCTD and NusG-NTD.

We extended these studies to full-length proteins step-by-step. In an initial experiment we added RNAP to ^{15}N-NusG, which led to a loss of almost all ^{15}N-NusG signals in the [^{1}H,^{15}N]-HSQC spectrum owing to the dramatic increase of the NusG rotational correlation time upon formation of the NusG:RNAP complex (Supplementary Figure S8A). Although NusG-NTD is only flexibly linked to NusG-CTD and NusG is supposed to interact with RNAP *via* NusG-NTD (4), NusG-CTD signals were not observable in the ^{15}N-NusG:RNAP complex. Thus, either NusG-CTD is sterically hindered in the complex so that it cannot move freely, or NusG-CTD interacts directly with RNAP. To exclude such direct NusG-CTD:RNAP interac-

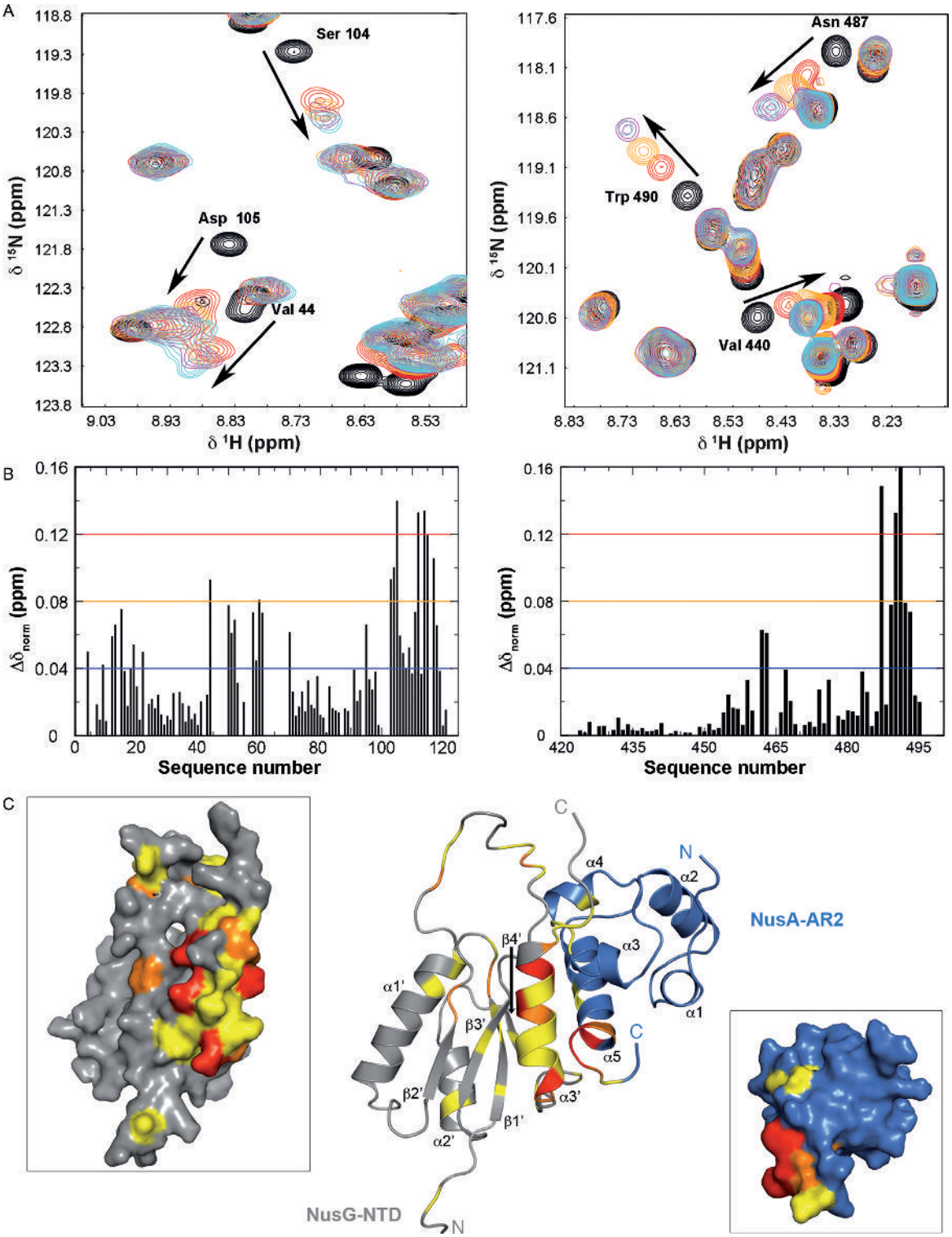


Figure 3. NusG-NTD:NusA-AR2 complex formation. (**A**, left) Sections of the [^{1}H,^{15}N]-HSQC-spectra of the titration of 140 μM ^{15}N-NusG-NTD with NusA-AR2. NusA-AR2 was added in a molar ratio of 1:0, black, 1:0.75, red, 1:1.25, orange, 1:2.5, magenta, and 1:3.5, cyan (stock concentration of NusA-AR2: 1.1 mM). (right) Sections of the [^{1}H,^{15}N]-HSQC-spectra of the titration of 100 μM ^{15}N-NusA-AR2 with NusG-NTD. Spectra corresponding to molar ratios 1:0, 1:0.5, 1:1, 1:2.5, and 1:3 are in black, red, orange, magenta, and cyan, respectively (stock concentration of NusG-NTD: 300 μM). Arrows indicate chemical shift changes during the titrations, selected signals are assigned. (**B**) HSQC-derived normalized chemical shift changes versus sequence position. (Left) $\Delta\delta_{norm}$ of ^{15}N-NusG-NTD on titration with NusA-AR2; (right) $\Delta\delta_{norm}$ of ^{15}N-NusA-AR2 on titration with NusG-NTD. Horizontal lines: significance levels of $\Delta\delta_{norm}$ (ppm) = 0.12, red; = 0.08, orange; = 0.04, blue. (**C**) Model of the NusA-AR2:NusG-NTD complex. The complex was generated with HADDOCK using the chemical shift perturbations of the [^{1}H,^{15}N]-HSQC titrations as restraints. The model with the best HADDOCK score is depicted. NusA-AR2 (PDB ID: 2K06), blue, and NusG-NTD (PDB ID: 1WCN), grey, are in cartoon representation. The normalized chemical shift changes from (**B**) are mapped on the structures (0.04 ppm < $\Delta\delta_{norm}$ < 0.08 ppm, yellow; 0.08 ppm < $\Delta\delta_{norm}$ < 0.12 ppm, orange; $\Delta\delta_{norm}$ > 0.12 ppm, red). Panels show the surface representations of NusG-NTD, left, and NusA-AR2, right.

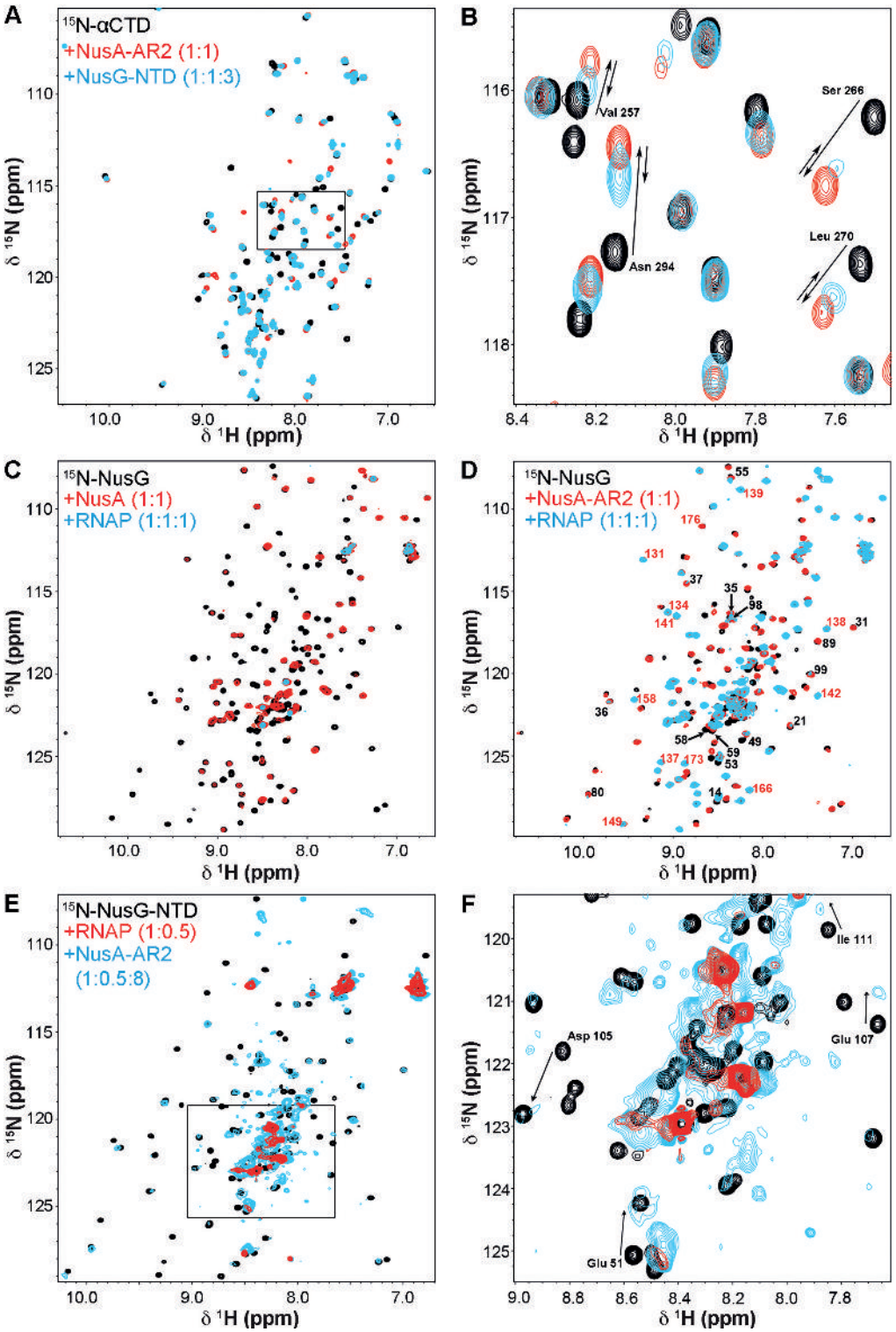


Figure 4. NusG-NTD:NusA-AR2 interaction in the presence of RNAP. (**A** and **B**) [^{1}H,^{15}N]-HSQC displacement experiment of NusA-AR2 from ^{15}N-αCTD by NusG-NTD. Black, ^{15}N-αCTD; red, ^{15}N-αCTD:NusA-AR2 = 1:1; blue, ^{15}N-αCTD:NusA-AR2:NusG-NTD = 1:1:3. The concentration of ^{15}N-αCTD was always 50 μM. The rectangle in (**A**) indicates the section as in (**B**). The arrows in (**B**) show the changes in the chemical shifts of selected residues. (**C**) NusG binds to NusA in the presence of RNAP. [^{1}H,^{15}N]-HSQC spectra of ^{15}N-NusG, black, ^{15}N-NusG in the presence of NusA (molar ratio 1:1), red, and ^{15}N-NusG in the presence of NusA and RNAP (molar ratio 1:1:1), cyan. (**D**) NusG binds to NusA-AR2 in the presence of RNAP. [^{1}H,^{15}N]-HSQC spectra of ^{15}N-NusG, black, ^{15}N-NusG in the presence of NusA-AR2 (molar ratio 1:1), red, and ^{15}N-NusG in the presence of NusA-AR2 and RNAP (molar ratio 1:1:1), cyan. Selected signals are labeled (black, NusG-NTD signals; red, NusG-CTD signals). The concentration of ^{15}N-NusG was 50 μM in all experiments in (**C**) and (**D**). (**E**) NusA-AR2 removes NusG-NTD from RNAP. [^{1}H,^{15}N]-HSQC spectra of ^{15}N-NusG-NTD, black, ^{15}N-NusG-NTD in the presence of RNAP (molar ratio 1:0.5), red, and ^{15}N-NusG-NTD in the presence of RNAP and NusA-AR2 (molar ratio 1:0.5:8), cyan. The concentration of ^{15}N-NusG was always 50 μM. The rectangle in (**E**) indicates the section as in (**F**). In (**F**), selected signals are assigned with arrows indicating changes in their chemical shifts corresponding to the complex formation of ^{15}N-NusG-NTD and NusA-AR2.

tion, we monitored isolated ^{15}N-NusG-CTD in the presence of RNAP in a separate experiment and found no changes in the [^{1}H,^{15}N]-HSQC spectra as compared to ^{15}N-NusG-CTD alone (Supplementary Figure S8B). Consequently, the loss of all ^{15}N-NusG signals upon RNAP addition solely originates from formation of the NusG-NTD:RNAP complex.

To probe the NusA:NusG interaction in the presence of full-length RNAP we added NusA to ^{15}N-NusG, leaving only NusG-CTD signals visible (Figure 4C and Supplementary Figure S8C). On addition of RNAP all signals decreased (Figure 4C and Supplementary Figure S8C). Thus, either (i) NusA:NusG interaction is disrupted by RNAP and both NusA and NusG bind individually to RNAP, or (ii) NusA:NusG remains intact and interacts with RNAP *via* NusA-NTD, or (iii) both. To eliminate interference by NusA-NTD:RNAP interactions, we repeated the experiment using isolated NusA-AR2 instead of full-length NusA (Figure 4D and Supplementary Figure S8D). When NusA-AR2 was present, the [^{1}H,^{15}N]-HSQC spectrum of ^{15}N-NusG showed chemical shift changes corresponding to NusG-NTD:NusA-AR2 complex formation (see also Figure 2D). On addition of RNAP the intensity of the NusG signals decreased, however, in contrast to the experiment with full-length NusA, both NusG-NTD and NusG-CTD signals remained visible with the chemical shifts of the NusG-NTD:NusA-AR2 complex. We conclude (i) that at least some of the NusG-NTD:NusA-AR2 complexes remain intact in the presence of RNAP and (ii) that these complexes cannot bind to RNAP in the absence of NusA-NTD. This confirms that NusA-AR2:NusG and NusG:RNAP formation are mutually exclusive. The decrease in signal intensity may be explained by dissociation of a certain portion of the NusG-NTD:NusA-AR2 complex so that NusG binds to RNAP, and NusA-AR2 either interacts with the αCTD of RNAP or remains free. Thus, with full length NusA, the NusG:NusA complex is stable and is connected to RNAP *via* NusA-NTD, although a fraction of NusG and NusA might interact with RNAP individually.

We next demonstrated that NusA-AR2 can remove NusG-NTD from RNAP (Figure 4E). As expected, ^{15}N-NusG-NTD signals were drastically diminished by addition of RNAP. However, they reappeared upon NusA-AR2 addition with the chemical shift perturbations typical for the NusG-NTD:NusA-AR2 complex. Hence, NusA-AR2 and RNAP compete for NusG-NTD.

NusA-AR2 induces pausing and blocks NusG suppression of pausing *in vitro*

We then asked if the NusA-AR2:NusG-NTD interaction affects transcription *in vitro*. For this we utilized a nucleic acid scaffold to generate a transcription elongation complex (TEC) which carries an 11 nt ^{32}P-labeled RNA primer base-paired to template DNA and flanked by non-template DNA (TEC11A, for details see Materials and Methods). Transcription was initiated by the addition of the four NTPs and stopped after 60 seconds. The TEC paused at several intrinsic pause sites in the template in the absence of additional transcription factors, with pause 1 being the most prominent (Figure 5; lane 2). Full-length NusG suppressed pause

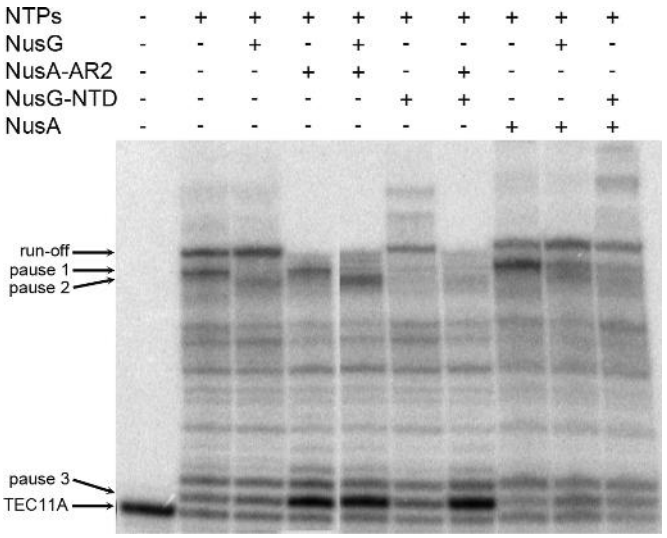


Figure 5. *In vitro* transcription assay for combinations of NusG, NusA, NusG-NTD and NusA-AR2. 23 % urea-polyacrylamide gel. The assembled TEC (TEC11A) was pre-incubated with NusG, NusA, NusA-AR2, NusG-NTD, combinations of these, or transcription buffer for 10 min at 25°C. Transcription was started by NTP addition and stopped after 60 s. TEC11A, run-off, and three pause sites are indicated by arrows.

1, increased run-off transcription and introduced a weak new pause, pause 2 (lane 3). NusA-AR2, interestingly, generated a strong pause at position 12C (pause 3), 1 nt downstream of the transcription start site, without influencing other pause sites (lane 4). Moreover, NusA-AR2 completely blocked run-off transcription (lane 4). When both proteins were present in equimolar amounts, pause 2 was enhanced, whereas pause 1 and run-off transcripts were suppressed (lane 5). Enhancement of NusG-dependent pause 2 is consistent with an interaction between NusG and NusA-AR2, possibly explained by the inability of NusG-NTD to enhance processivity when bound to NusA-AR2. The NusA-AR2-dependent pause 3 was not influenced by NusG (lane 5). NusG-NTD yielded similar results as full-length NusG. It suppressed pause 1 (lane 6), and this suppression was abrogated by NusA-AR2 (lane 7). As with full-length NusG, NusG-NTD did not affect NusA-AR2-induced pause 3 (lane 7).

Full-length NusA enhanced pause 1, but did not, however, induce pausing at pause 3 (lane 8). This suggests that the NusA-AR2 domain in full-length NusA was still bound to the NusA-SKK domain, and was not free to interact with the initiating TEC. Unlike NusA-AR2, suppression of pause 1 by full-length NusG or NusG-NTD was not completely abrogated by full-length NusA, possibly because NusA-AR2 remains bound to the NusA-SKK domain during elongation, and is not available to interact with NusG-NTD (lanes 8–10).

The ability of isolated NusA-AR2 to pause the TEC at 12C (pause 3) was unexpected. It suggests that at least early in elongation, when RNA has not yet extruded from the exit channel, NusA-AR2 may still be bound to the SKK domain and may thus be unavailable to interact with αCTD. In addition, the ability of NusA-AR2 to induce a strong pause

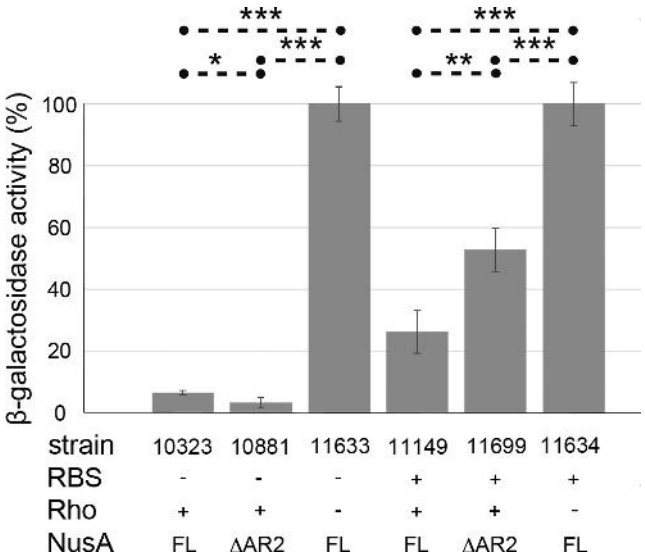


Figure 6. Deletion of NusA-AR2 affects termination at λ*tR1*. β-galactosidase reporter assays were performed with *lacZ* fusions λ*cI857 – pR – cro(ΔRBS) – nutR – tR1 – cII::lacZ* and λ*cI857 – pR – cro*$_{27}$ *– nutR – tR1 – cII::lacZ*. Strains are derivatives of *E. coli* MDS42. Cells were assayed for β-galactosidase activity (Miller units) after overnight growth at 37°C. β-galactosidase activity of cells with defective Rho was set to 100% (strains 11633 and 11634). *P* values are < 0.05 (*), < 0.01 (**), or < 0.001 (***). RBS +/−, functional/defective RBS; Rho +/−, functional/defective Rho; FL, full length.

implies that NusA-AR2 might act as a regulatory element during elongation if dissociated from αCTD.

NusA-ΔAR2 blocks Rho-dependent termination at λ*tR1 in vivo*

NusA suppresses termination at certain Rho-dependent sites, e.g. within λ*tR1* (46). We propose that the NusA:NusG interaction contributes to this effect. To support this hypothesis we asked if a deletion of NusA-AR2 (NusA-ΔAR2) affected termination at λ*tR1 in vivo* (Figure 6). We performed β-galactosidase assays using two fusions to measure termination: λ*cI857 – pR – cro(ΔRBS) – nutR – tR1 – cII::lacZ* and λ*cI857 – pR – cro*$_{27}$ *– nutR – tR1 – cII::lacZ*. Termination at λ*tR1* is indicated by low β-galactosidase activity. Strains carrying a mutation in the *rho* gene show no termination at λ*tR1*; β-galactosidase activity of these strains was thus set to 100%. The efficiency of termination at λ*tR1* was 93% when *cro* translation was prevented by an RBS mutation (strain 10323), and reduced to 74 % when *cro* was translated (strain 11149). Ribosomes reduce the amount of free RNA upstream to λ*tR1* that is available to Rho, and thus block a Rho-binding site (*rut*) in *cro* (47).

We found that termination efficiency was significantly impaired (50%) in the *nusA-ΔAR2* mutant only when *cro* mRNA was translated (compare strain 10881 to strain 10323 and strain 11699 to strain 11149). Our results suggest that NusA-ΔAR2 may compete with Rho binding near λ*nutR*, the only *rut* site available when the *cro* transcript is occluded by ribosomes. When *cro* is not translated, Rho can attach to the free *cro rut* site (48). Competition with Rho by NusA-ΔAR2 can be explained by constitutive binding of NusA-ΔAR2 to RNA *via* NusA-SKK. Consistent with the *in vitro* studies described above, this implies that full length NusA may still be, at least partially, in the autoinhibited state and unable to bind *rut* RNA at λ*tR1*. It also raises the possibility that NusG may activate RNA binding of full-length NusA by displacing NusA-AR2 from the NusA-SKK domain. Further experimentation will be needed to address these questions.

Possible regulatory functions of the NusG-NTD:NusA-AR2 interaction

Transcription factors NusA and NusG act independently to slow or accelerate transcription elongation, respectively. The two factors can also function synergistically in restoring termination by the partially defective RhoE134K mutant (49). Here, however, we demonstrate a direct interaction between NusA and NusG. Formation of a complex between NusG-NTD and NusA-AR2 may explain their various *in vivo* and *in vitro* combinatorial regulatory effects (Figure 7).

1. NusA-AR2 supports NusG recruitment.
   NusG:NusA interaction may be involved in recruiting NusG to the TEC. ChIP-chip data suggest that NusA and NusG associate with RNAP when the TEC has exited the promoter region, with NusG attaching after NusA does (50). The delay in NusG association may be due to a competition between σ^{70} region 2 and NusG-NTD, since both bind to the RNAP β'CH (7,51–53). Recall that the σ^{70} region 2 can remain bound to the TEC even when σ^{70} region 4 has dissociated from the β flap, allowing NusA-NTD binding (54,55). After promoter escape, NusA attaches to the αCTD *via* NusA-AR2, to the β flap *via* NusA-NTD, and to nascent RNA *via* NusA-SKK (10,15,18,56). Thus, stable tethering of NusA to the TEC may not require continuous binding of NusA-AR2 to αCTD. NusA-AR2 could, therefore, bind NusG-NTD without disrupting the NusA:TEC complex. In this model, NusA serves as a long linker to increase the local concentration of NusG, facilitating NusG recruitment to the TEC and displacement of σ^{70} region 2 from the β'CH at the transition from initiation to elongation.
2. NusG-NTD:NusA-AR2 interaction assists transcription:translation coupling.
   As a second function, we suggest that NusG:NusA interaction plays a role in coupling transcription and translation. NusG connects these two processes by physically linking RNAP and the leading ribosome *via* NusG-NTD:RNAP and NusG-CTD:S10 interactions (4,8). The NusA:NusG interaction could serve to resynchronize transcription and translation by coordinating the movements of RNAP and the ribosome. If translation is slowed, transcription could likewise be slowed by the temporary removal of NusG-NTD from the TEC by NusA-AR2. Since the NusG:ribosome connection remains intact, transcription and translation can be kinetically resynchronized. Also, the initial coupling of transcription and translation may occur *via* the NusA:NusG

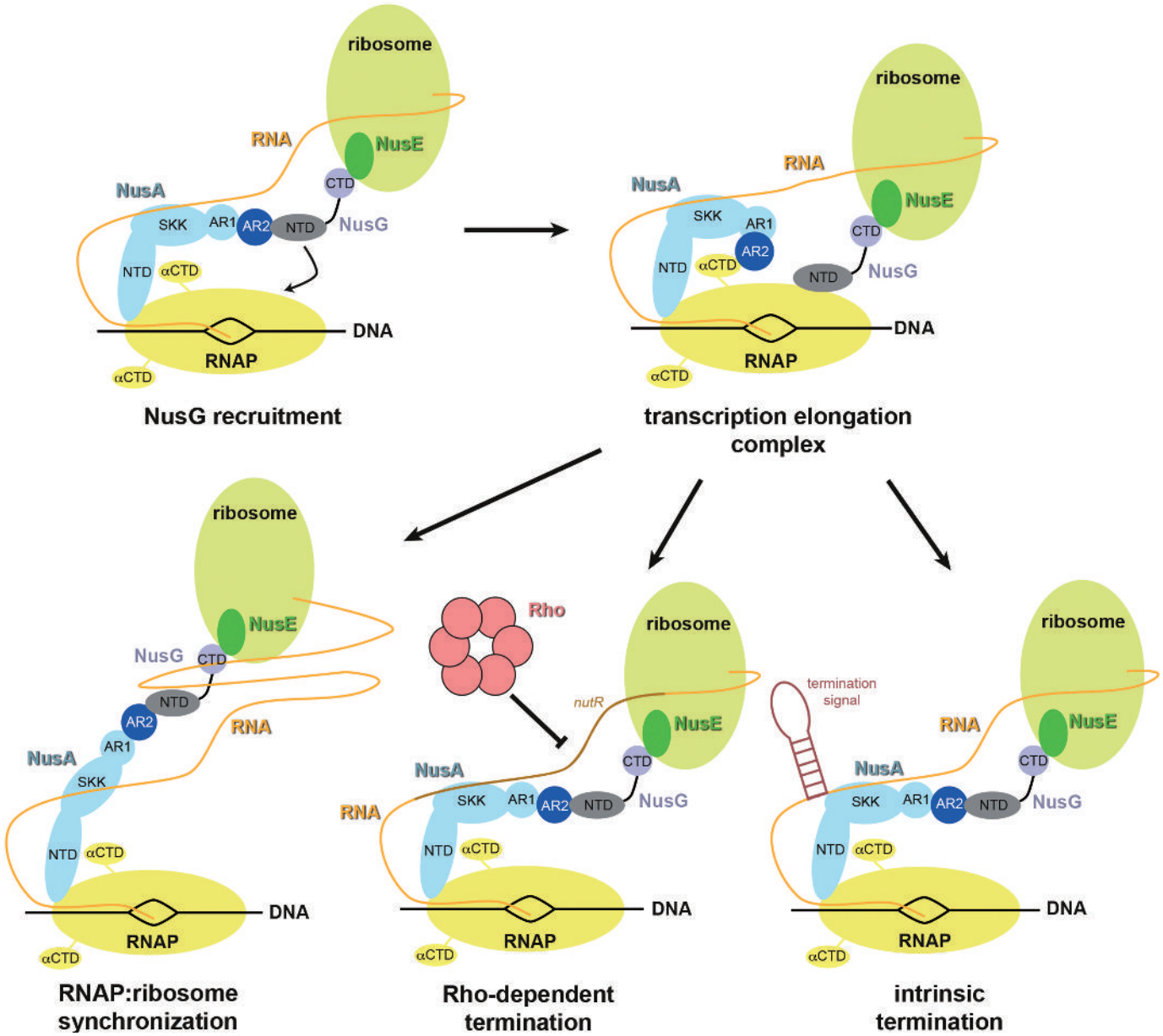


Figure 7. Possible functions of NusG:NusA interaction in transcription regulation. First, the NusA:NusG interaction may play a role in the recruitment of NusG to the TEC. Second, it may provide a long linker between RNAP and the ribosome consisting of NusA and NusG, which would allow resynchronization of transcription:translation coupling. Third, the NusG:NusA interaction may release the autoinhibition of NusA allowing constitutive binding of NusA-SKK to RNA, so that recruitment sites for Rho are blocked and Rho-dependent termination is decreased. Forth, NusA-AR2 may abstract NusG-NTD at intrinsic termination sites facilitating the release of nucleic acids and enhancing intrinsic termination.

linker. This would explain the apparent late association of NusG with the TEC, as detected by ChIP-chip experiments (50).

3. NusG-NTD:NusA-AR2 interaction regulates transcription termination.
   The NusA:NusG interaction could influence context-dependent intrinsic or Rho-dependent transcription termination. In the former, the TEC pauses at an intrinsic termination signal, enters an elemental pause state, and then isomerizes into the termination state where the termination hairpin is formed (57). NusA-AR2 can remove NusG-NTD from RNAP, resulting in loss of both NusA-AR2 and NusG-NTD contacts to the TEC (Figure 4E). Loss of these interactions might partially destabilize the TEC, open the clamp around the nucleic acids, and facilitate intrinsic termination. In contrast, the NusA-AR2:NusG-NTD interaction might decrease Rho-dependent termination. Binding of NusA-AR2 to NusG-NTD would release NusA autoinhibition, enhancing binding of NusA-SKK to nascent RNA to block Rho recruitment.

The K_D values for the various interactions suggest that scenario 1 is the most probable. This scenario is also con-

sistent with a global survey of distribution of transcription factors (50). The relevance of NusG-NTD:NusA-AR2 interaction in detail will need further experimental clarification, but the finding that NusG interacts directly with NusA may explain the various effects of these transcription factors on elongation and termination reported here and earlier.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

M.S., S.H.K., P.R. thank Ramona Heißmann for excellent technical support. We also thank Dr J. Drögemüller for carefully reading the manuscript and Drs J. Drögemüller and R. Washburn for helpful discussions.

FUNDING

Deutsche Forschungsgemeinschaft [Ro 617/21-1 to P.R.]; Ludwig-Schaefer-Scholarship 2015 from the Columbia University Medical Center (to P.R.); National Institutes of Health [GM37219 to M.G.]. Funding for open access charge: Universität Bayreuth.
Conflict of interest statement. None declared.

REFERENCES

1. Werner,F. and Grohmann,D. (2011) Evolution of multisubunit RNA polymerases in the three domains of life. *Nat. Rev. Microbiol.*, **9**, 85–98.
2. Mooney,R.A., Artsimovitch,I. and Landick,R. (1998) Information processing by RNA polymerase: recognition of regulatory signals during RNA chain elongation. *J. Bacteriol.*, **180**, 3265–3275.
3. Burns,C.M., Richardson,L.V. and Richardson,J.P. (1998) Combinatorial effects of NusA and NusG on transcription elongation and rho-dependent termination in *Escherichia coli*. *J. Mol. Biol.*, **278**, 307–316.
4. Mooney,R.A., Schweimer,K., Rösch,P., Gottesman,M.E. and Landick,R. (2009) Two structurally independent domains of *E. coli* NusG create regulatory plasticity *via* distinct interactions with RNA polymerase and regulators. *J. Mol. Biol.*, **391**, 341–358.
5. Werner,F. (2012) A nexus for gene expression-molecular mechanisms of Spt5 and NusG in the three domains of life. *J. Mol. Biol.*, **417**, 13–27.
6. Sevostyanova,A., Belogurov,G.A., Mooney,R.A., Landick,R. and Artsimovitch,I. (2011) The β subunit gate loop is required for RNA polymerase modification by RfaH and NusG. *Mol Cell*, **43**, 253–262.
7. Martinez-Rucobo,F.W., Sainsbury,S., Cheung,A.C. and Cramer,P. (2011) Architecture of the RNA polymerase-Spt4/5 complex and basis of universal transcription processivity. *EMBO J.*, **30**, 1302–1310.
8. Burmann,B.M., Schweimer,K., Luo,X., Wahl,M.C., Stitt,B.L., Gottesman,M.E. and Rösch,P. (2010) A NusE:NusG complex links transcription and translation. *Science*, **328**, 501–504.
9. Friedman,D.I., Schauer,A.T., Baumann,M.R., Baron,L.S. and Adhya,S.L. (1981) Evidence that ribosomal protein S10 participates in control of transcription termination. *Proc. Natl. Acad. Sci. U.S.A.*, **78**, 1115–1118.
10. Yang,X., Molimau,S., Doherty,G.P., Johnston,E.B., Marles-Wright,J., Rothnagel,R., Hankamer,B., Lewis,R.J. and Lewis,P.J. (2009) The structure of bacterial RNA polymerase in complex with the essential transcription elongation factor NusA. *EMBO Rep.*, **10**, 997–1002.
11. Ha,K.S., Toulokhonov,I., Vassylyev,D.G. and Landick,R. (2010) The NusA N-terminal domain is necessary and sufficient for enhancement of transcriptional pausing *via* interaction with the RNA exit channel of RNA polymerase. *J. Mol. Biol.*, **401**, 708–725.
12. Ma,C., Mobli,M., Yang,X., Keller,A.N., King,G.F. and Lewis,P.J. (2015) RNA polymerase-induced remodelling of NusA produces a pause enhancement complex. *Nucl. Acids Res.*, **43**, 2829–2840.
13. Eisenmann,A., Schwarz,S., Prasch,S., Schweimer,K. and Rösch,P. (2005) The *E. coli* NusA carboxy-terminal domains are structurally similar and show specific RNAP- and λN interaction. *Protein Sci.*, **14**, 2018–2029.
14. Worbs,M., Bourenkov,G.P., Bartunik,H.D., Huber,R. and Wahl,M.C. (2001) An extended RNA binding surface through arrayed S1 and KH domains in transcription factor NusA. *Mol. Cell*, **7**, 1177–1189.
15. Mah,T.F., Li,J., Davidson,A.R. and Greenblatt,J. (1999) Functional importance of regions in escherichia coli elongation factor NusA that interact with RNA polymerase, the bacteriophage lambda N protein and RNA. *Mol. Microbiol.*, **34**, 523–537.
16. Prasch,S., Schwarz,S., Eisenmann,A., Wöhrl,B.M., Schweimer,K. and Rösch,P. (2006) Interaction of the intrinsically unstructured phage lambda N protein with *E. coli* NusA. *Biochemistry*, **45**, 4542–4549.
17. Mishra,S., Mohan,S., Godavarthi,S. and Sen,R. (2013) The interaction surface of a bacterial transcription elongation factor required for complex formation with an antiterminator during transcription antitermination. *J. Biol. Chem.*, **288**, 28089–28103.
18. Schweimer,K., Prasch,S., Santhanam,S.P., Bubunenko,M., Gottesman,M.E. and Rösch,P. (2011) NusA interaction with the α-subunit of *E. coli* RNA polymerase is *via* the UP-element site and releases autoinhibition. *Structure*, **19**, 945–954.
19. Artsimovitch,I. and 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.
20. Cardinale,C.J., Washburn,R.S., Tadigotla,V.R., Brown,L.M., Gottesman,M.E. and Nudler,E. (2008) Termination factor Rho and its cofactors NusA and NusG silence foreign DNA in *E. coli*. *Science*, **320**, 935–938.
21. Mason,S.W., Li,J. and Greenblatt,J. (1992) Host factor requirements for processive antitermination of transcription and suppression of pausing by the N protein of bacteriophage lambda. *J. Biol. Chem.*, **267**, 19418–19426.
22. Shankar,S., Hatoum,A. and Roberts,J.W. (2007) A transcription antiterminator constructs a NusA-dependent shield to the emerging transcript. *Mol. Cell*, **27**, 914–927.
23. Weisberg,R.A. and Gottesman,M.E. (1999) Processive antitermination. *J. Bacteriol.*, **181**, 359–367.
24. Borukhov,S., Lee,J. and Laptenko,O. (2005) Bacterial transcription elongation factors: New insights into molecular mechanism of action. *Mol. Microbiol.*, **55**, 1315–1324.
25. Roberts,J.W., Shankar,S. and Filter,J.J. (2008) RNA polymerase elongation factors. *Annu. Rev. Microbiol.*, **62**, 211–233.
26. Li,J., Mason,S.W. and Greenblatt,J. (1993) Elongation factor NusG interacts with termination factor Rho to regulate termination and antitermination of transcription. *Genes Dev.*, **7**, 161–172.
27. Nehrke,K.W., Zalatan,F. and Platt,T. (1993) NusG alters Rho-dependent termination of transcription *invitro* independent of kinetic coupling. *Gene Expr.*, **3**, 119–133.
28. Kohno,T., Kusunoki,H., Sato,K. and Wakamatsu,K. (1998) A new general method for the biosynthesis of stable isotope-enriched peptides using a decahistidine-tagged ubiquitin fusion system: An application to the production of mastoparan-X uniformly enriched with ^{15}N and ^{15}N/^{13}C. *J. Biomol. NMR*, **12**, 109–121.
29. Burmann,B.M., Schweimer,K., Scheckenhofer,U. and Rösch,P. (2011) Domain interactions of the transcription:Translation coupling factor *E.coli* NusG are intermolecular and transient. *Biochem. J.*, **435**, 783–789.
30. Drögemüller,J., Strauss,M., Schweimer,K., Wöhrl,B.M., Knauer,S.H. and Rösch,P. (2015) Exploring RNA polymerase regulation by NMR spectroscopy. *Sci. Rep.*, **5**, 10825–10835.
31. 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.
32. Eisenmann,A., Schwarz,S., Prasch,S., Schweimer,K. and Rosch,P. (2005) The *E. coli* NusA carboxy-terminal domains are structurally similar and show specific RNAP- and lambdaN interaction. *Protein Sci.*, **14**, 2018–2029.

33. Eisenmann,A., Schwarz,S., Rösch,P. and Schweimer,K. (2004) Sequence-specific ^{1}H, ^{13}C, ^{15}N resonance assignments and secondary structure of the carboxyterminal domain of the *E. coli* transcription factor NusA. *J. Biomol. NMR*, **28**, 193–194.
34. Artsimovitch,I., Svetlov,V., Murakami,K.S. and Landick,R. (2003) Co-overexpression of *Escherichia coli* RNA polymerase subunits allows isolation and analysis of mutant enzymes lacking lineage-specific sequence insertions. *J. Biol. Chem.*, **278**, 12344–12355.
35. Kashlev,M., Martin,E., Polyakov,A., Severinov,K., Nikiforov,V. and Goldfarb,A. (1993) Histidine-tagged RNA polymerase: Dissection of the transcription cycle using immobilized enzyme. *Gene*, **130**, 9–14.
36. Sambrook,J., Fritsch,E.F. and Maniatis,T. (1994) *Molecular Cloning – A Laboratory Manual*. Cold Spring Harbor Laboratory Press, NY.
37. Meyer,O. and Schlegel,H.G. (1983) Biology of aerobic carbon monoxide-oxidizing bacteria. *Annu. Rev. Microbiol.*, **37**, 277–310.
38. Vitiello,C.L., Kireeva,M.L., Lubkowska,L., Kashlev,M. and Gottesman,M. (2014) Coliphage HK022 nun protein inhibits RNA polymerase translocation. *Proc. Natl. Acad. Sci. U.S.A.*, **111**, E2368–E2375.
39. Johnson,B.A. (2004) Using NMRView to visualize and analyze the NMR spectra of macromolecules. *Methods Mol. Biol.*, **278**, 313–352.
40. de Vries,S.J., van Dijk,M. and Bonvin,A.M. (2010) The HADDOCK web server for data-driven biomolecular docking. *Nat. Protoc.*, **5**, 883–897.
41. Posfai,G., Plunkett,G. 3rd, Feher,T., Frisch,D., Keil,G.M., Umenhoffer,K., Kolisnychenko,V., Stahl,B., Sharma,S.S., de Arruda,M. *et al.* (2006) Emergent properties of reduced-genome *Escherichia coli*. *Science*, **312**, 1044–1046.
42. Svenningsen,S.L., Costantino,N., Court,D.L. and Adhya,S. (2005) On the role of Cro in lambda prophage induction. *Proc. Natl. Acad. Sci. U.S.A.*, **102**, 4465–4469.
43. Schrödinger,L. (2010) *The PyMOL Molecular Graphics System, version 1.3*. Schrödinger, LLC, Mannheim.
44. Sievers,F., Wilm,A., Dineen,D., Gibson,T.J., Karplus,K., Li,W., Lopez,R., McWilliam,H., Remmert,M., Soding,J. *et al.* (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using clustal omega. *Mol. Syst. Biol.*, **7**, 539.
45. Drögemüller,J., Strauß,M., Schweimer,K., Jurk,M., Rösch,P. and Knauer,S.H. (2015) Determination of RNA polymerase binding surfaces of transcription factors by NMR spectroscopy. *Sci. Rep.*, **5**, 16428–16441.
46. Lau,L.F., Roberts,J.W. and Wu,R. (1983) RNA polymerase pausing and transcript release at the lambda *tR1* terminator *in vitro*. *J. Biol. Chem.*, **258**, 9391–9397.
47. Faus,I. and Richardson,J.P. (1990) Structural and functional properties of the segments of lambda *cro* mRNA that interact with transcription termination factor Rho. *J. Mol. Biol.*, **212**, 53–66.
48. Graham,J.E. and Richardson,J.P. (1998) *Rut* sites in the nascent transcript mediate Rho-dependent transcription termination *in vivo*. *J. Biol. Chem.*, **273**, 20764–20769.
49. Muteeb,G., Dey,D., Mishra,S. and Sen,R. (2012) A multipronged strategy of an anti-terminator protein to overcome Rho-dependent transcription termination. *Nucleic Acids Res.*, **40**, 11213–11228.
50. Mooney,R.A., Davis,S.E., Peters,J.M., Rowland,J.L., Ansari,A.Z. and Landick,R. (2009) Regulator trafficking on bacterial transcription units *in vivo*. *Mol. Cell*, **33**, 97–108.
51. Belogurov,G.A., Vassylyeva,M.N., Svetlov,V., Klyuyev,S., Grishin,N.V., Vassylyev,D.G. and Artsimovitch,I. (2007) Structural basis for converting a general transcription factor into an operon-specific virulence regulator. *Mol. Cell*, **26**, 117–129.
52. Arthur,T.M. and Burgess,R.R. (1998) Localization of a sigma70 binding site on the N terminus of the *Escherichia coli* RNA polymerase beta' subunit. *J. Biol. Chem.*, **273**, 31381–31387.
53. Young,B.A., Anthony,L.C., Gruber,T.M., Arthur,T.M., Heyduk,E., Lu,C.Z., Sharp,M.M., Heyduk,T., Burgess,R.R. and Gross,C.A. (2001) A coiled-coil from the RNA polymerase beta' subunit allosterically induces selective nontemplate strand binding by sigma(70). *Cell*, **105**, 935–944.
54. Bar-Nahum,G. and Nudler,E. (2001) Isolation and characterization of sigma(70)-retaining transcription elongation complexes from escherichia coli. *Cell*, **106**, 443–451.
55. Kuznedelov,K., Minakhin,L., Niedziela-Majka,A., Dove,S.L., Rogulja,D., Nickels,B.E., Hochschild,A., Heyduk,T. and Severinov,K. (2002) A role for interaction of the RNA polymerase flap domain with the sigma subunit in promoter recognition. *Science*, **295**, 855–857.
56. Mah,T.F., Kuznedelov,K., Mushegian,A., Severinov,K. and Greenblatt,J. (2000) The alpha subunit of *E. coli* RNA polymerase activates RNA binding by NusA. *Genes Dev.*, **14**, 2664–2675.
57. Weixlbaumer,A., Leon,K., Landick,R. and Darst,S.A. (2013) Structural basis of transcriptional pausing in bacteria. *Cell*, **152**, 431–441.

SUPPLEMENTARY INFORMATION

for

Transcription is regulated by NusA:NusG interaction

Martin Strauß[1], Christal Vitiello[2], Kristian Schweimer[1], Max Gottesman[2,3], Paul Rösch[1], and Stefan H. Knauer[1,*]

[1]Lehrstuhl Biopolymere und Forschungszentrum für Bio-Makromoleküle, Universität Bayreuth, 95447 Bayreuth, Germany

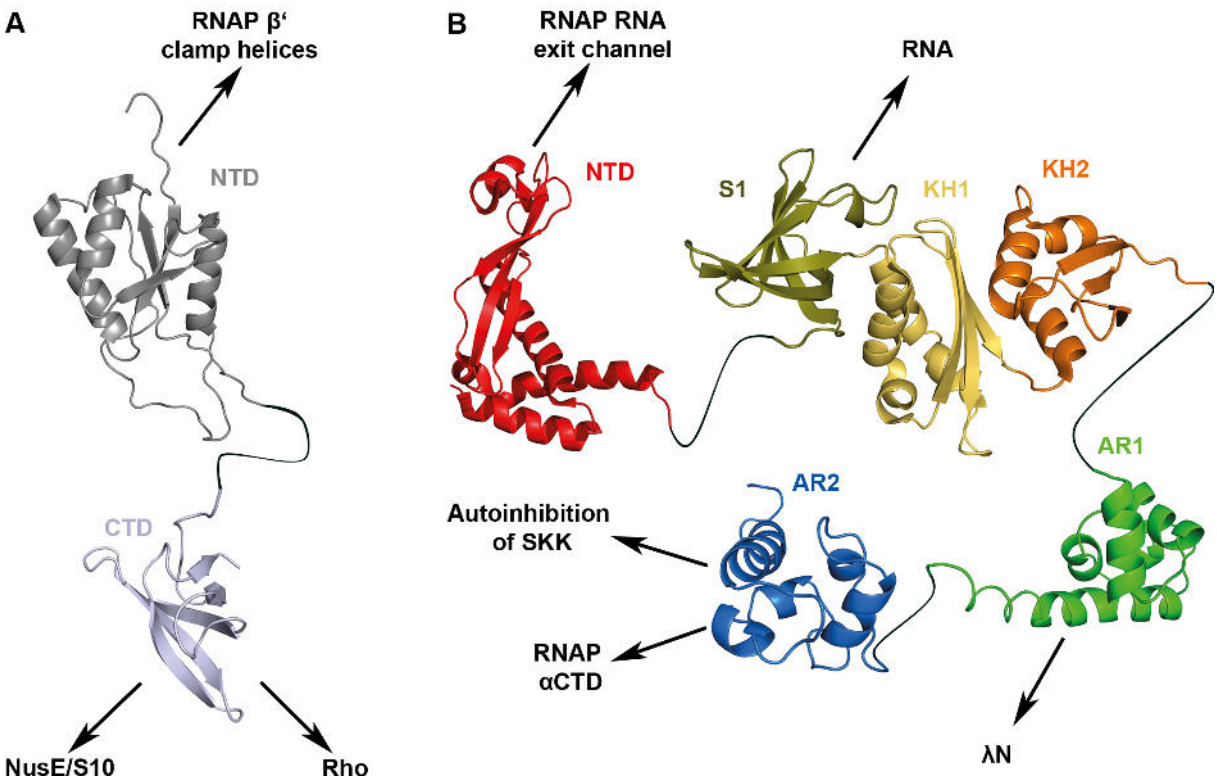
[2]Department of Microbiology and Immunology, Columbia University Medical Center, New York, NY 10032, USA

[3]Department of Biochemistry and Molecular Biophysics, Columbia University Medical Center, New York, NY 10032, USA

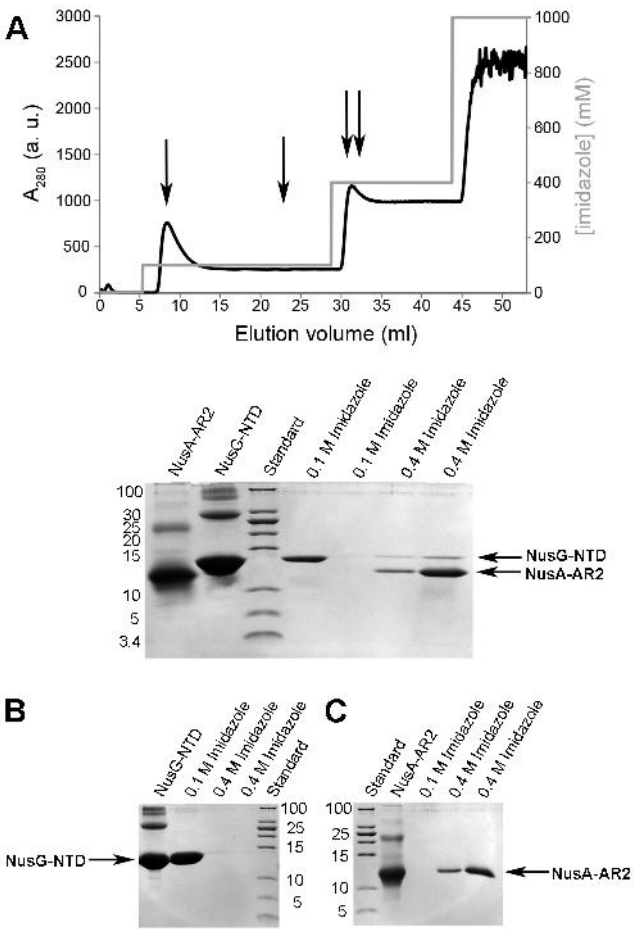
* To whom correspondence should be addressed. Tel: 0049 921 553868; Fax: 0049 921 16490459; Email: stefan.knauer@uni-bayreuth.de

Present Address: Martin Strauß, Department of Microbiology and Immunology, Columbia University Medical Center, New York, NY 10032, USA Department, Institution, Town, State, Postcode, Country

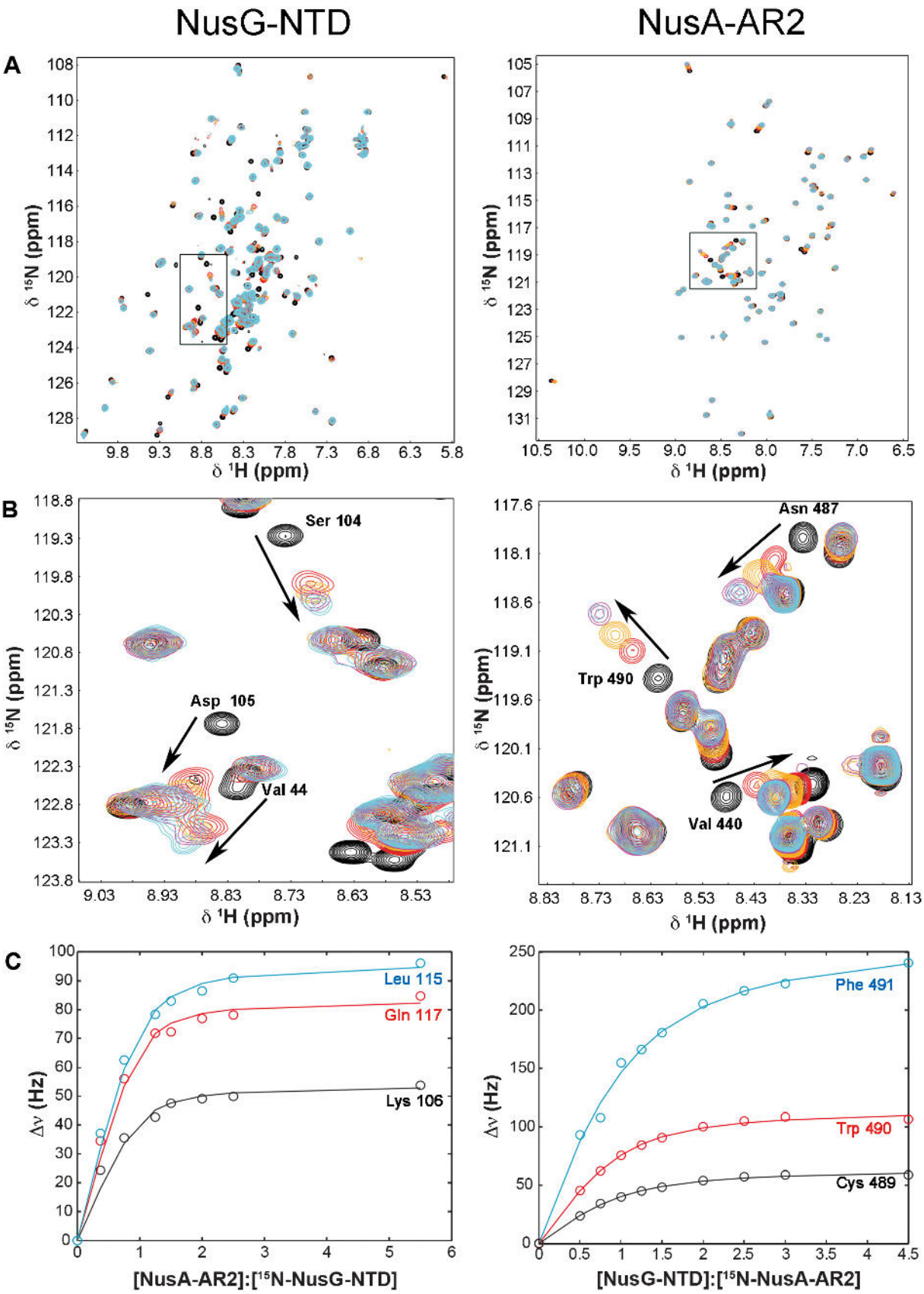
Contents:



Supplementary Figure S1. Three-dimensional structures of (A) NusG and (B) NusA. Protein structures are in cartoon representation. Arrows indicate the interaction partner(s) of individual domains. (**A**) NusG from *E. coli.* NusG-NTD, gray, PDB ID: 2K06; NusG-CTD, light blue, PDB ID: 2JVV; flexible linker, black line. (**B**) NusA from *E. coli.* NusA-NTD, red, PDB ID: 2KWP; NusA-S1, olive; NusA-KH1, yellow; NusA-KH2, orange (as no structure of *E. coli* NusA-SKK is available the structure of *Thermotoga maritima* NusA-SKK is shown, PDB ID: 1HH2); NusA-AR1, green, PDB ID: 1WCL; NusA-AR2, blue, PDB ID: 1WCN; linker, black line.

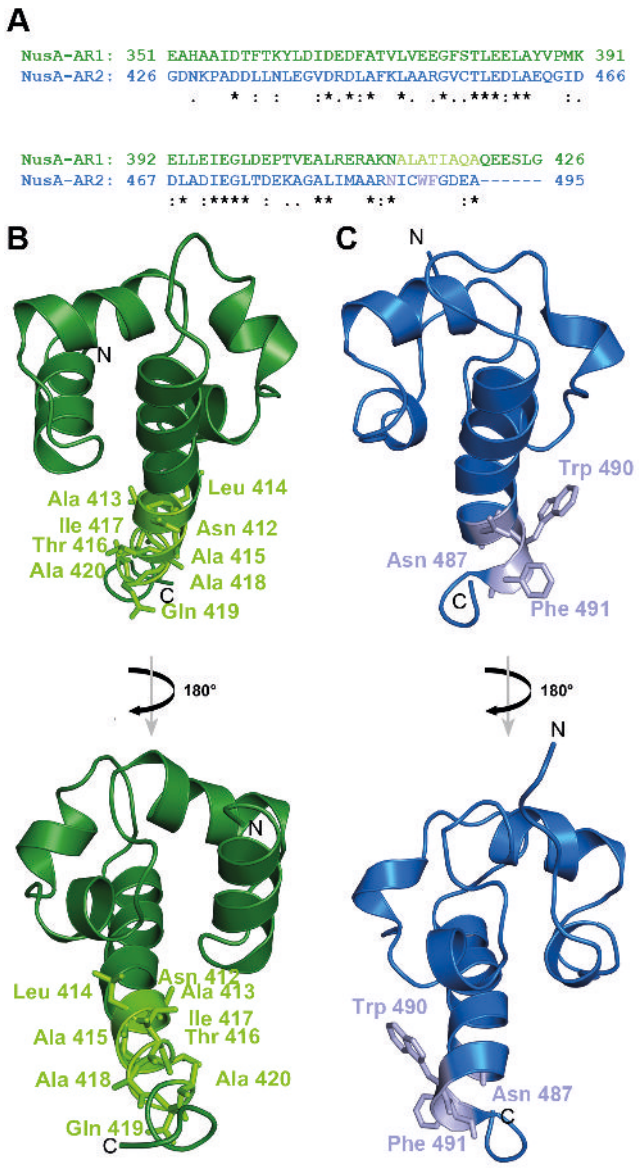


Supplementary Figure S2. Pull-down of NusG-NTD with His_{10}-NusA-AR2. (**A**) His_{10}-NusA-AR2 (200 μM) and NusG-NTD (400 μM) were preincubated for 15 min and then applied to a 1 ml HisTrap column. After washing, stepwise elution was carried out with 100 and 400 mM imidazole. (Upper panel) Chromatogram of the pull-down assay. Arrows indicate the fractions analyzed by sodium dodecylsulfate (SDS) polyacrylamide gel electrophoresis. (lower panel) 20 % SDS polyacrylamide gel of samples taken during the pull-down assay. NusA-AR2, pure His_{10}-NusA-AR2; NusG-NTD, pure NusG-NTD; 0.1 M imidazole, elution with 100 mM imidazole; 0.4 M imidazole, elution with 400 mM imidazole. (**B,C**) Control experiments with (**B**) NusG-NTD and (**C**) His_{10}-NusA-AR2. Isolated NusG-NTD (400 μM) or His_{10}-NusA-AR2 (200 μM) was applied to the column and treated like in (**A**).

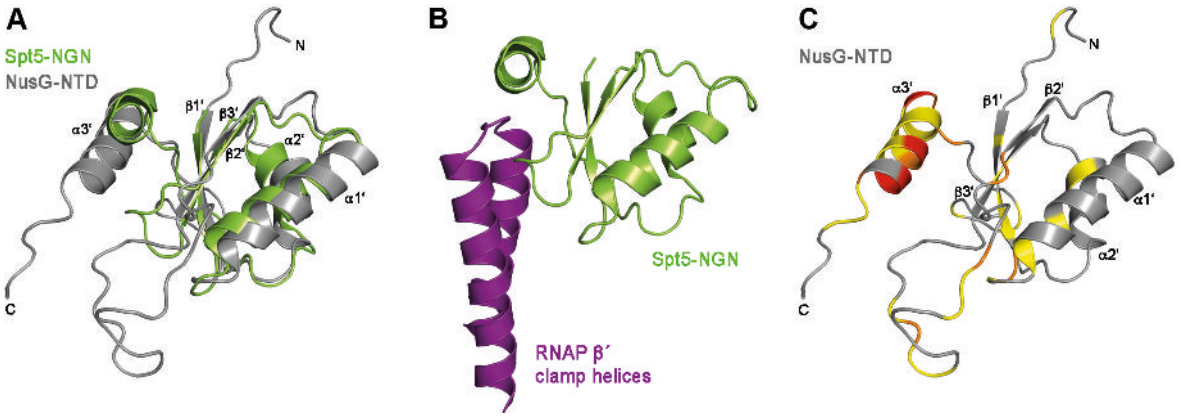


Supplementary Figure S3. Determination of K_D values of the NusA-AR2:NusG-NTD complex. (**A**, left) [^{1}H,^{15}N]-HSQC titration of ^{15}N-NusG-NTD (140 μM) with NusA-AR2. NusA-AR2 was added in

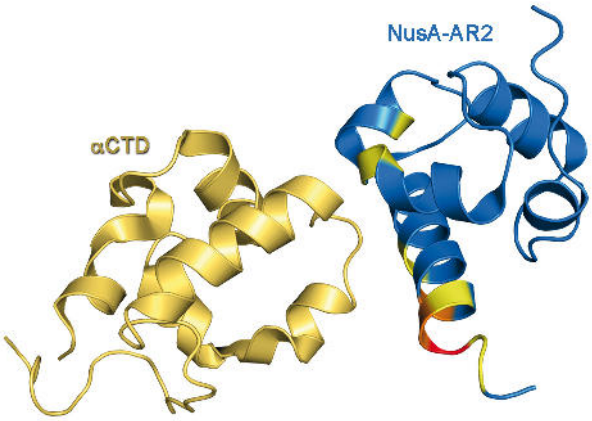
molar ratios of 1:0, black, 1:0.75, red, 1:1.25, orange, 1:2.5, magenta, and 1:3.5, cyan. (right) [^{1}H,^{15}N]-HSQC titration of ^{15}N-NusA-AR2 with NusG-NTD. Spectra corresponding to molar ratios 1:0, 1:0.5, 1:1, 1:2.5, and 1:3 are in black, red, orange, magenta, and cyan, respectively. (**B**) Magnifications of (**A**). Selected signals are labeled. (**C**) Backbone amide chemical shift perturbations for selected residues obtained form (**A**) *vs.* molar ratio of the titration partners. (Left) ^{15}N-NusG-NTD+NusA-AR2; (right) ^{15}N-NusA-AR2 + NusG-NTD. The lines represent nonlinear least squares best fits of the normalized changes in the ^{1}H and ^{15}N chemical shifts, based on a bimolecular equilibrium binding model. The optimized average K_D values are 13 μM for ^{15}N-NusG-NTD + NusA-AR2 and 35 μM for ^{15}N-NusA-AR2 + NusG-NTD, yielding an overall K_D of approximately 22 μM for the NusA-AR2:NusG-NTD interaction.



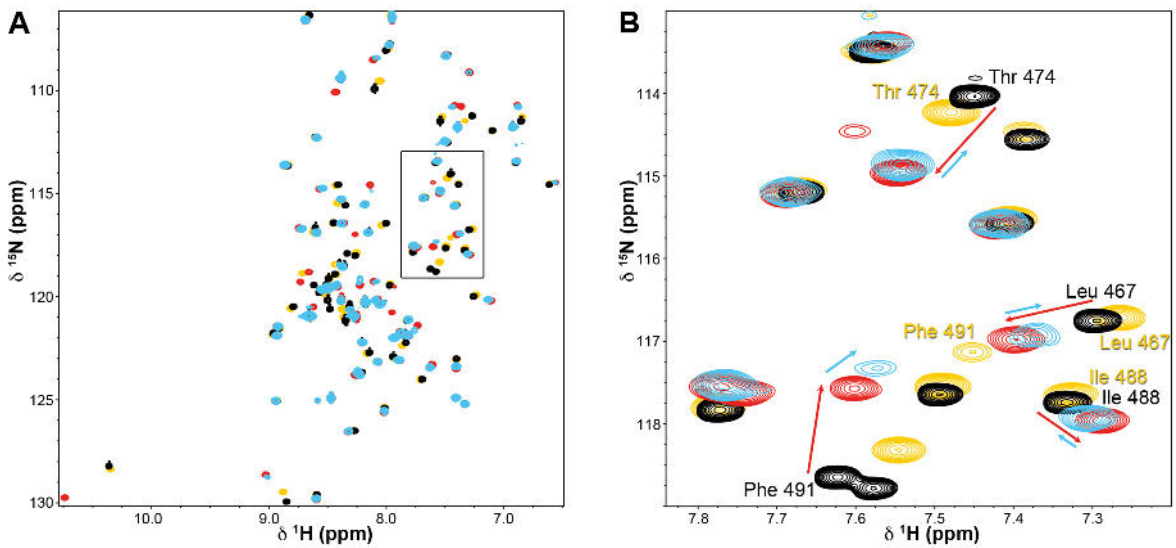
Supplementary Figure S4. Comparison of NusA-AR1 and NusA-AR2. (**A**) Amino acid sequence alignment of NusA-AR1 and NusA-AR2. Asterisk, identical amino acids; colon, conservation between groups of strongly similar properties; dot, conservation between groups of weakly similar properties. (**B,C**) Structures of (**A**) NusA-AR1, green, and (**B**) NusA-AR2, blue, both in cartoon representation. Residues of NusA-AR2 which are strongly affected by NusG-NTD binding ($\Delta\delta_{norm}$ > 0.12 ppm) as well as corresponding residues in NusA-AR1 are shown as sticks in light colours and labelled. PDB IDs: NusA-AR1, 1WCL; NusA-AR2, 1WCN. .



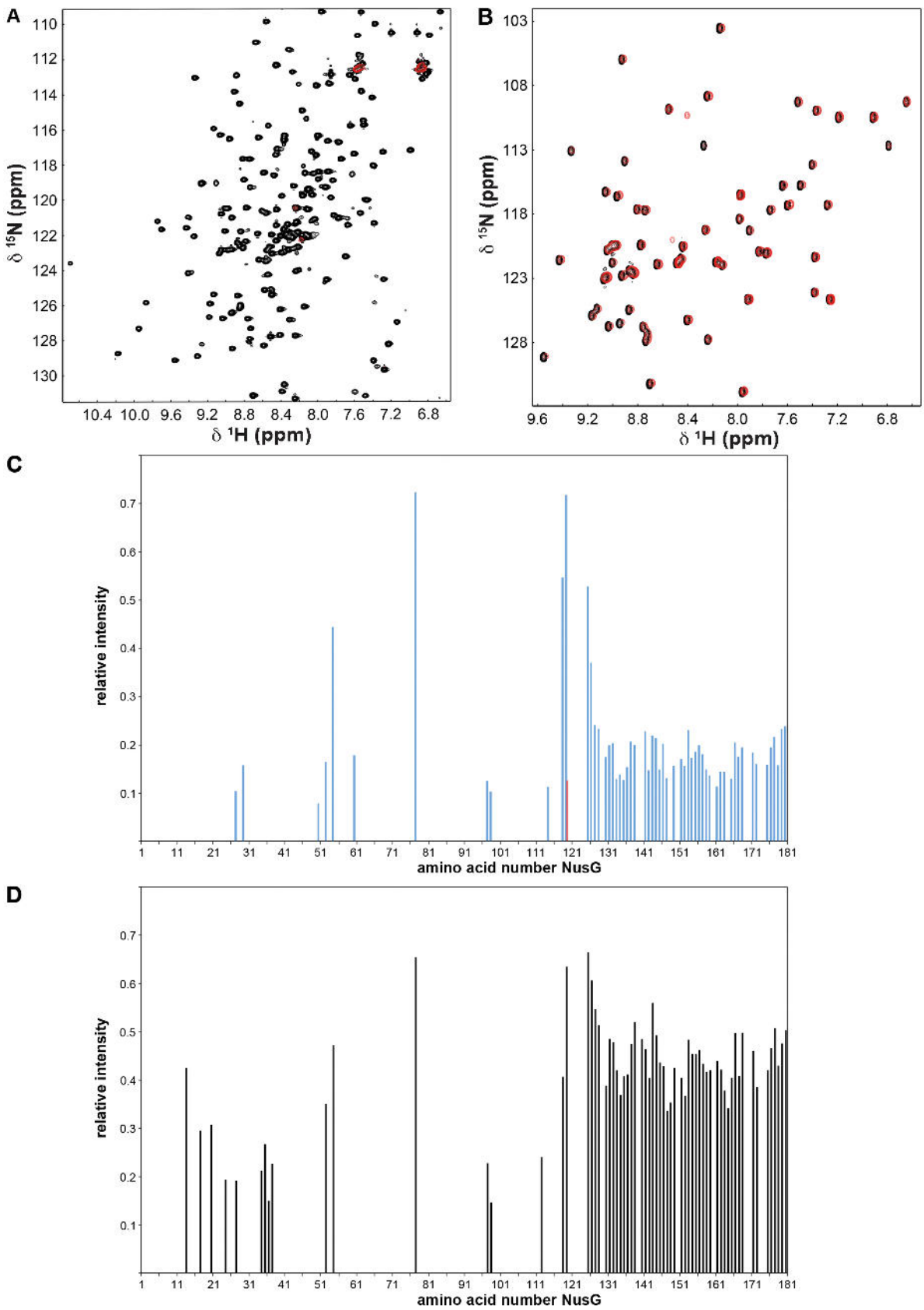
Supplementary Figure S5. Binding of NusG proteins to RNAP β'CH. (**A**) Superposition of Spt5-NusG N-terminal domain (NGN) from *Pyrococcus furiosus* (*P. furiosus*, green, PDB ID: 3QQC) and NusG-NTD from *E. coli* (grey, PDB ID: 2K06), both in cartoon representation. (**B**) Spt5-NGN bound to the β'CH (purple) in *P. furiosus* (PDB ID: 3QQC). (**C**) NusG-NTD in the same orientation as in (**A**). Residues that are affected by the interaction with NusA-AR2 are in red (strongly affected), orange (moderately affected), and yellow (slightly affected), see Figure 3.



Supplementary Figure S6. The NusA-AR2 binding sites for αCTD and NusG-NTD overlap. Solution structure of the NusA-AR2:αCTD complex (PDB ID: 2JZB, cartoon representation). Dark yellow, αCTD; blue, NusA-AR2. Residues of NusA-AR2 that are affected by the interaction with NusG-NTD are in red (strongly affected), orange (moderately affected), and yellow (slightly affected), see Figure 3.



Supplementary Figure S7. NusG-NTD:NusA-AR2 interaction in the presence of RNAP. (**A**) [^{1}H,^{15}N]-HSQC displacement experiment of ^{15}N-NusA-AR2 from αCTD by NusG-NTD. Black, ^{15}N-NusA-AR2 (100 μM); red, ^{15}N-NusA-AR2:αCTD = 1:1 (100 μM each); blue, ^{15}N-NusA-AR2:αCTD:NusG-NTD = 1:1:5; yellow, ^{15}N-NusA-AR2:NusG-NTD = 1:3. (**B**) Detail of (**A**). Red arrows, chemical shift changes of ^{15}N-NusA-AR2 upon ^{15}N-NusA-AR2:αCTD complex formation; blue arrows, chemical shift changes of ^{15}N-NusA-AR2 upon addition of NusG-NTD.



Supplementary Figure S8. NusG:RNAP *vs.* NusG:NusA *vs.* NusA:NusG. (**A**) [^{1}H,^{15}N]-HSQC spectrum of 50 μM ^{15}N-NusG in the absence, black, or presence, red, of RNAP in equimolar

concentration. (**B**) [^{1}H,^{15}N]-HSQC spectrum of 50 μM ^{15}N-NusG-CTD in the absence, black, or presence, red, of RNAP in equimolar concentration. (**C**) NusG binds to NusA in the presence of RNAP. Intensity plots of the titration of Fig. 4C. Relative intensities were calculated in respect to free NusG. Blue, ^{15}N-NusG:NusA = 1:1; red, ^{15}N-NusG:NusA:RNAP = 1:1:1 (**D**) NusG:NusA-AR2 complex remains intact in the presence of RNAP. Intensity plot of the displacement experiment in Fig. 4D. Relative intensities were calculated as ratio of intensities of ^{15}N-NusG signals in the presence of NusA-AR2 and RNAP (1:1:1) and ^{15}N-NusG signals in the presence of NusA-AR2 (1:1).

7.3 Einzelarbeit C

Johanna Drögemüller*, Martin Strauß*, Kristian Schweimer, Birgitta M. Wöhrl, Stefan H. Knauer, Paul Rösch (2015): **Exploring RNA polymerase regulation by NMR Spectroscopy**, *Scientific Reports,* Juni 2015 4;5: 10825.

* Beide Autoren haben in gleichem Maße zur Arbeit beigetragen.

SCIENTIFIC REPORTS

OPEN

Exploring RNA polymerase regulation by NMR spectroscopy

Johanna Drögemüller*, Martin Strauß*, Kristian Schweimer, Birgitta M. Wöhrl, Stefan H. Knauer & Paul Rösch

Received: 13 February 2015
Accepted: 22 April 2015
Published: 04 June 2015

RNA synthesis is a central process in all organisms, with RNA polymerase (RNAP) as the key enzyme. Multisubunit RNAPs are evolutionary related and are tightly regulated by a multitude of transcription factors. Although *Escherichia coli* RNAP has been studied extensively, only little information is available about its dynamics and transient interactions. This information, however, are crucial for the complete understanding of transcription regulation in atomic detail. To study RNAP by NMR spectroscopy we developed a highly efficient procedure for the assembly of active RNAP from separately expressed subunits that allows specific labeling of the individual constituents. We recorded [^{1}H,^{13}C] correlation spectra of isoleucine, leucine, and valine methyl groups of complete RNAP and the separately labeled β′ subunit within reconstituted RNAP. We further produced all RNAP subunits individually, established experiments to determine which RNAP subunit a certain regulator binds to, and identified the β subunit to bind NusE.

The synthesis of RNA is a central process in cells that is carried out by DNA-dependent RNA polymerases (RNAPs). All cellular genomes are transcribed by multisubunit RNAPs that are evolutionary related. In spite of their differences in size and complexity, RNAPs share overall architecture, active-site organization, mechanism of catalysis, and the principles of interactions with nucleic acids[1].

In bacteria, the RNAP core enzyme consists of five subunits, 2xα, β, β', and ω, with different structural and functional roles[2,3]. The C-terminal domains (CTD) of the α subunits (αCTD) are target for many regulatory proteins and are thus key factors for the regulation of transcription[4,5]. Dimerization of the N-terminal domains (NTD) of the α subunits initiates the RNAP assembly process[6]. Next, the β subunit attaches to the α dimer, followed by recruitment of the β' and the ω subunit[6,7]. While the β and β' subunits constitute the active center of RNAP, the ω subunit plays a structural rather than a functional role as it is supposed to bind to the N- and C-termini of the β' subunit to support its proper folding as well as the assembly of β'ω with the $\alpha_2\beta$ complex[7,8]. The σ factor binds to RNAP at the initiation of transcription to form holo RNAP. σ is essential for the recognition and melting of promoter regions, and it leaves RNAP in later stages of transcription[9,10].

Initiation, elongation, and termination of transcription are highly regulated by transcription factors that bind to the transcription elongation complex (TEC) and modify the RNAP[11]. NusG, for example, enhances the transcription rate and suppresses pausing[12]. It interacts with the RNAP β' clamp helices (β'CH) and the RNAP β gate loop (βGL)[13,14]. In contrast to NusG, NusA modifies RNAP to induce pausing and to modulate intrinsic as well as Rho-dependent termination of transcription (reviewed in [15,16]). NusA, NusG, NusB, and NusE can combine with the TEC and certain RNA sequences to form an antitermination complex which is able to read through termination signals, a process that is essential for efficient transcription of ribosomal DNA or the DNA of lambdoid phages[17]. While NusG-NTD mediates RNAP binding, NusG-CTD interacts with NusE in the NusE:NusB complex[18,19]. As NusE, also known as ribosomal protein S10, can be part of the 30S subunit of the ribosome[20], NusG physically links RNAP and the ribosome, thus coupling transcription and translation[18]. Moreover, NusE may also directly interact with RNAP[21].

Lehrstuhl Biopolymere und Forschungszentrum für Bio-Makromoleküle, Universität Bayreuth, Universitätsstraße 30, 95447 Bayreuth, Germany. *These authors contributed equally to this work. Correspondence and requests for materials should be addressed to S.H.K. (email: stefan.knauer@uni-bayreuth.de)

Numerous crystallographic studies on prokaryotic and eukaryotic RNAPs have elucidated the structural basis of RNAP architecture and gave insights into its function (reviewed in[22]). However, RNAP regulation is heavily dependent on intra- and intermolecular dynamics as well as transient interactions with regulators, which are difficult to study in atomic detail by X-ray crystallography or electron microscopy.

Although nuclear magnetic resonance (NMR) spectroscopy of supramolecular complexes is aggravated by ^{1}H-^{1}H and ^{1}H-^{13}C dipolar interactions that lead to fast relaxation of the magnetization and therefore loss of signal intensity, deuteration[23], application of more sophisticated pulse sequences like transverse relaxation optimized spectroscopy (TROSY), and use of [^{1}H,^{13}C] methyl group probes result in improvements of spectral quality so that proteins up to 670 kDa have been studied successfully[24–26].

Encouraged by these results, we improved the assembly of *E. coli* RNAP from its separately expressed subunits and started to study this reconstituted RNAP by NMR spectroscopy. We use [^{1}H,^{13}C] correlation spectra of isoleucine, leucine, and valine methyl groups in complete RNAP and in the β' subunit of reassembled RNAP to study transcription regulator interactions with RNAP, and we propose to extend this method to other RNAP subunits and RNAPs of other organisms.

Results and Discussion

***In vitro* RNAP assembly, purification, and biochemical characterization.** Bacterial RNAP without ω subunit, but containing σ factor, can be reconstituted from individually expressed and separately purified protein subunits[27–29]. Analysis of elongating RNAP requires, however, inclusion of the ω subunit and omission of the σ factor. Hence, we combined the cell pellets containing the individually expressed subunits α, β, β', or ω, respectively, in lysis buffer with 8 M urea. After cell lysis the lysate was stirred for one hour and subsequently urea was removed by stepwise dialysis. The assembled core RNAP was purified by Ni^{2+} affinity chromatography, and RNAP eluted from a size exclusion chromatography (SEC) column in peaks at 47.5 ml and 54.8 ml (Fig. 1a), corresponding to molecular masses of 980 kDa and 507 kDa, respectively. Analysis of the peak fractions by sodium dodecylsulfate polyacrylamide gel electrophoresis (SDS-PAGE) clearly showed that both peaks contained all RNAP subunits, although the 980 kDa fractions included a high amount of impurities (Fig. 1b). The calculated molecular mass of RNAP of 390 kDa suggests that the protein from the second peak is correctly reconstituted RNAP free of major contaminants. As a reference, we used RNAP assembled *in vivo* ($RNAP^{native}$), where the genes of the subunits were located on a single plasmid. Indeed, $RNAP^{native}$ eluted from the SEC column in a main peak coinciding with the 507 kDa peak (Fig. 1a). An activity assay testing the ability of RNAP to elongate an RNA primer showed that protein from the 507 kDa peak and $RNAP^{native}$ were both functionally identical (Fig. 1c). Therefore, we refer to active reassembled RNAP as $RNAP^{active}$ in contrast to inactive reassembled RNAP ($RNAP^{inactive}$) from the 980 kDa peak.

The far-UV circular dichroism (CD) spectra of $RNAP^{native}$ and $RNAP^{active}$ are very similar (Fig. 1d), with the typical characteristics of a folded protein. In contrast, the spectrum of $RNAP^{inactive}$ is of lower intensity with less distinct minima, in particular the minimum characteristic for α-helical elements at 208 nm is nearly absent, indicating that $RNAP^{active}$ and $RNAP^{native}$ are folded similarly, whereas $RNAP^{inactive}$ is at least partially unfolded or misfolded.

$RNAP^{active}$ was reapplied onto a SEC column to analyze if it was in equilibrium with $RNAP^{inactive}$. The enzyme eluted in a single peak at the same volume as before, indicating that the protein is stable on the time scale of these experiments (Fig. 1a). Additionally, we could increase the yield of correctly assembled RNAP by de- and renaturation of $RNAP^{inactive}$. Subsequent SEC again yielded peaks at 46.3 and 55.5 ml corresponding to the two RNAP states (Fig. 1a). Hence, at least a portion of the misassembled RNAP could be reconstituted into $RNAP^{active}$.

Overall, the yield was 30–60 mg of $RNAP^{active}$ per liter of bacterial cultures producing α, β, β', and ω, the purity exceeding 95%, similar to the published protocols for RNAP assembly lacking ω. Although the ω subunit of RNAP, encoded by the *rpoZ* gene, is neither essential for cell viability nor for RNAP function, the activity of RNAP lacking σ increases when reassembled in the presence of ω[29–31]. In *rpoZ* deletion strains RNAP copurifies with GroEL and loses its activity upon GroEL removal. However, activity can be regained by denaturation and renaturation of RNAP in the presence of ω[31]. ω was suggested to have important functions in folding of the β' subunit, in preventing β' from aggregation as well as in promoting the assembly of $\alpha_2\beta$ with β'ω[7]. Thus, its presence during reconstitution might reduce the amount of misfolded or misassembled RNAP.

Overall, this assembly and purification strategy allows efficient production of complete, pure, and active core RNAP from separately expressed subunits. In contrast to earlier protocols, purification of one or all individual subunits prior to RNAP assembly is unnecessary, the ω subunit is part of the assembled RNAP, and the presence of the initiation factor σ is not required, so that purified RNAP can be used directly in an elongation context. Finally, by using SEC as final purification step we selectively purify active RNAP and exclude all misassembled and inactive variants, a step that was omitted in most previous protocols.

Purification of individual RNAP subunits and analysis of their secondary structure. We expressed and purified all RNAP subunits separately (α, β, β', and ω) with high yield and purity of > 95%, allowing structural analyses (Supplementary Fig. 1). Additionally, the ββ' complex was assembled from individually expressed subunits and purified according to the protocol used for the assembly of

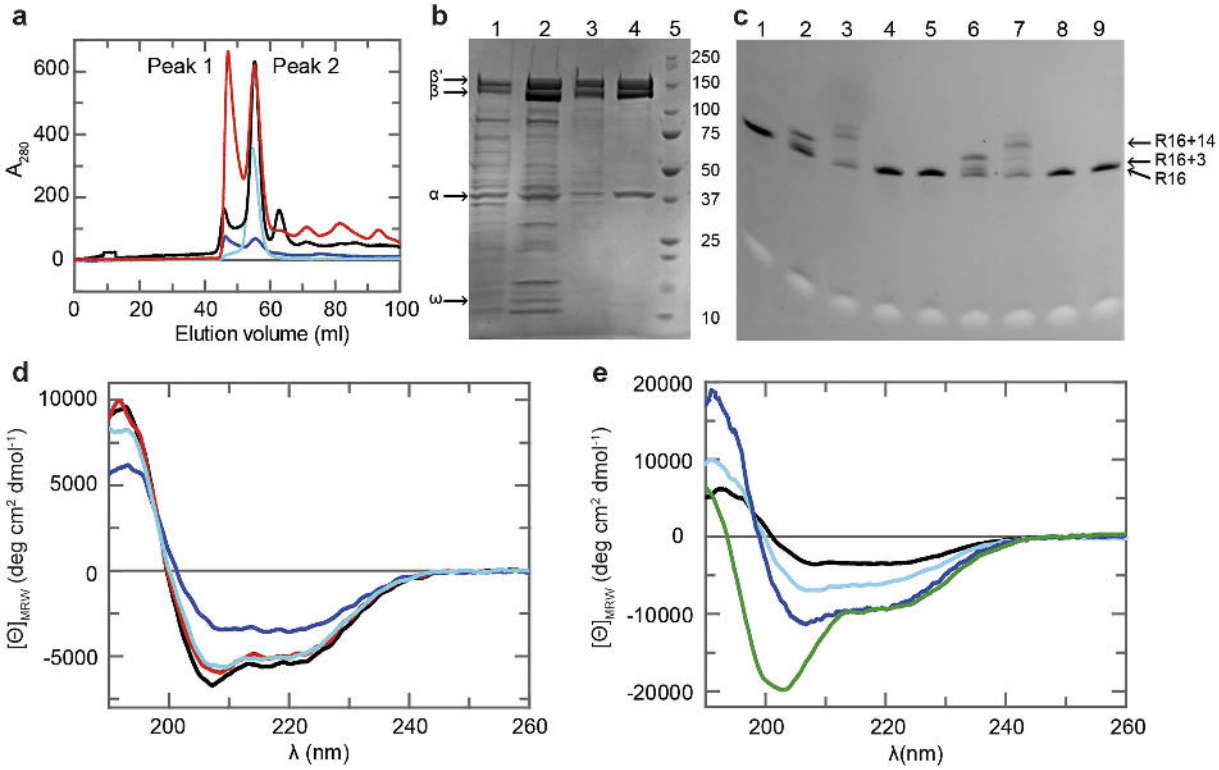


Figure 1. Purification of *in vitro* assembled RNAP. (**a**) Gel filtration chromatograms from an S200 column. Red: combined fractions after Ni^{2+} affinity chromatography; cyan: $RNAP^{active}$; blue: $RNAP^{inactive}$ after de- and renaturation; black: $RNAP^{native}$ (**b**) 4-20% gradient SDS-polyacrylamide gel (Roti-Page, Carl Roth, Karlsruhe, Germany) of aliquots taken during RNAP purification after staining with Coomassie Blue. In lanes 1-4 2 μg protein were applied. Soluble fraction of the assembled RNAP after dialysis (lane 1); combined fractions after Ni^{2+} affinity chromatography (lane 2); SEC peak 1 (lane 3); SEC peak 2 (lane 4); Precision Plus Protein Standard (BioRad, Munich, Germany, lane 5). (**c**) RNAP activity assay, 20% SDS-polyacrylamide gel. 3 pmol RNA were loaded in each lane. Either ATP and CTP or ATP, CTP and GTP were added allowing extension of a 16mer RNA (R16) by 3 or 14 nt, respectively. The arrows indicate the migration positions of R16 and the elongated RNAs. R16, untreated (lane 1); $RNAP^{native}$, elongation by 3 nt (lane 2) or 14 nt (lane 3); $RNAP^{inactive}$, elongation by 3 nt (lane 4) or 14 nt (lane 5); $RNAP^{active}$, elongation by 3 nt (lane 6) or 14 nt (lane 7); control reaction without RNAP, elongation by 3 nt (lane 8) or 14 nt (lane 9). (**d**) Far-UV CD-spectra of 0.6 μM $RNAP^{native}$, black; 0.6 μM $RNAP^{inactive}$, blue; 0.5 μM $RNAP^{active}$, red; 0.6 μM ββ' complex, cyan. (**e**) Far-UV CD-spectra of the separately expressed and purified RNAP subunits. 2.5 μM α, blue; 0.6 μM β, cyan; 1.1 μM β', black; 10 μM ω, green.

RNAP. All proteins were soluble, and although β was isolated from inclusion bodies it showed no tendency to precipitate up to concentrations of 120 μM after refolding. In contrast to previous publications, our protocol yielded soluble β'[28,32].

The far-UV CD spectra of α, β, and β' show the typical characteristics of structured proteins (Fig. 1e), and although the CD spectrum of the ω subunit exhibits the least distinct features, ω does not appear to be completely unfolded. Indeed, ω possesses a structured NTD, followed by an unstructured C-terminus[7] which is in agreement with the [^{1}H,^{15}N]-heteronuclear single quantum coherence (HSQC) spectrum of ^{15}N-labeled ω that shows very low signal dispersion (Supplementary Fig. 2), indicating that the isolated ω is only very poorly folded and might adopt its final structure only upon binding to β' or the complete RNAP. Subunits β and β' represent the largest part of RNAP and the CD spectrum of the ββ' complex is indeed nearly identical to that of $RNAP^{native}$ (Fig. 1d), suggesting that the isolated ββ' complex is assembled as it is in $RNAP^{native}$.

NusG-NTD interacts with β and β' while NusA-NTD binds to β and NusA-AR2 to α. As no activity assay can be conducted for the individual RNAP subunits, their integrity was checked by testing their ability to interact with transcription factors NusG and NusA whose RNAP binding sites are known. NusG consists of two domains that are flexibly connected[19]. It enhances RNAP processivity and reduces pausing by binding to RNAP *via* its NTD[12]. Thus, we first asked which RNAP subunit is the target site

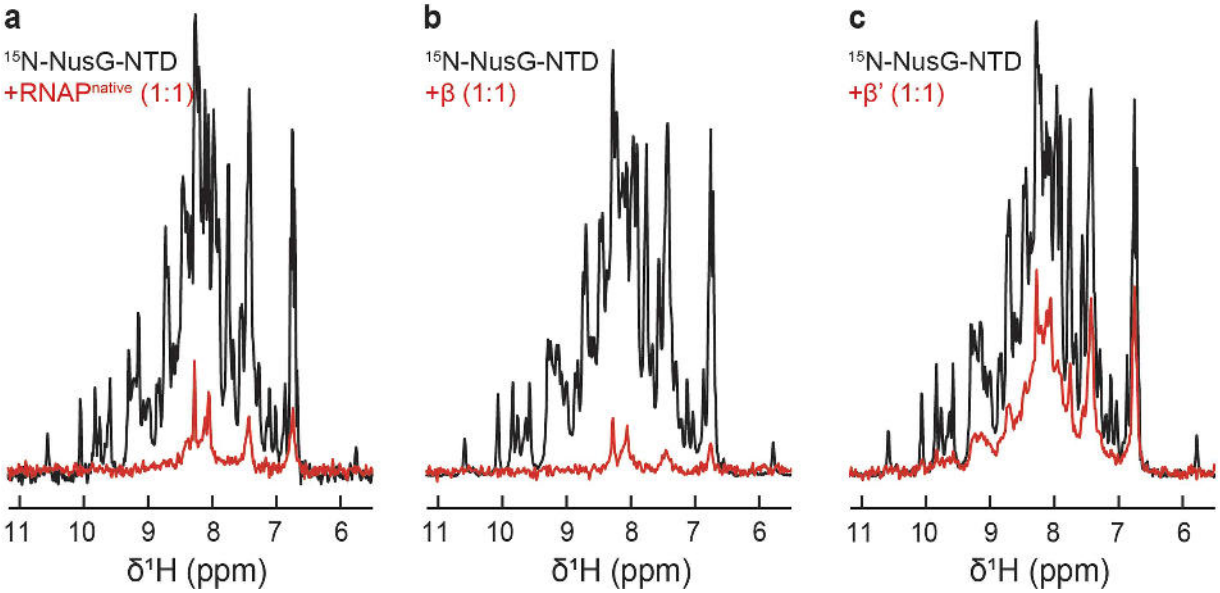


Figure 2. NusG-NTD interaction with RNAP, β, and β'. 1D [^{1}H,^{15}N]-HSQC spectra of the amide region of 30 μM ^{15}N-NusG-NTD in the absence, black, and in the presence of equimolar concentrations, red, of (**a**) RNAP^{native}, (**b**) β subunit, or (**c**) β' subunit.

for NusG-NTD. Upon addition of RNAP^{native}, the signals of ^{15}N-NusG-NTD in the one dimensional (1D) [^{1}H,^{15}N]-HSQC spectrum disappeared, except for a few signals in the random coil area as the resonances of ^{15}N-NusG-NTD are broadened significantly by the dramatic increase in the rotation correlation time due to the formation of the NusG-NTD:RNAP^{native} complex (Fig. 2a). Similarly, addition of isolated β or β' to ^{15}N-NusG-NTD lead to the loss of ^{15}N-NusG-NTD signal intensity (Fig. 2b,c). In contrast, the spectrum remained unaltered upon addition of α or ω (Supplementary Fig. 3), clearly demonstrating that NusG-NTD interacts only with β and β'. When β' was added, the loss of signal intensity was not as dramatic as it was upon addition of RNAP^{native} or β. This can be attributed to inaccuracies in concentration or to a lower affinity of NusG-NTD to β' as compared to complete RNAP or β. Our results are in good agreement with the known binding sites of NusG-NTD, i.e. the β'CH and the βGL (Fig. 3)[13,14].

During transcription, NusA decreases the elongation rate of RNAP, induces pausing, modulates intrinsic and Rho-dependent termination, and is part of the antitermination complex (reviewed in[15,16]). *E. coli* NusA consists of six domains, an NTD, three RNA binding domains (S1, KH1, KH2) that together form the SKK domain, and two C-terminal acidic repeat domains, AR1 and AR2[33,34]. While the interaction partner of NusA-AR2 is the αCTD of RNAP, NusA-NTD binds to the β flap-tip helix[5,35,36].

The NusA-NTD interaction with RNAP and its subunits was probed as with NusG-NTD. The disappearance of ^{15}N-NusA-NTD signals in the presence of RNAP^{native} confirms complex formation (Fig. 4a). However, addition of either β or β' led to an only slight decrease of ^{15}N-NusA-NTD signals, even in the presence of a twofold molar excess of the RNAP subunit (Fig. 4b,c), the effect being slightly more pronounced for the β subunit. In contrast, the signal decrease was more severe when the ββ' complex was added (Fig. 4d). To address the question whether this effect was due to a higher binding affinity or because of the increase in the molecular mass, we determined the observed amide proton transverse relaxation rate R_2 (R_2^{obs}) of free NusA-NTD and of NusA-NTD after addition of β, β', or ββ' in equimolar amounts by spin-echo experiments. R_2^{obs} of NusA-NTD increased in the presence of the individual subunits and the ββ' complex (R_2^{obs}: NusA-NTD, 50 s^{-1}; NusA-NTD+β, 130 s^{-1}; NusA-NTD+β', 90 s^{-1}; NusA-NTD+ββ', 190 s^{-1}). Assuming that R_2^{obs} is population-averaged, the fraction of unbound NusA-NTD was calculated according to equation (3). While the actual R_2 of NusA-NTD corresponds to its R_2^{obs} value, the R_2 values of NusA-NTD completely bound to β, β' or ββ' were estimated based on the proportionality of R_2 and the molecular mass. When β or ββ' were present, approximately 80% of NusA-NTD molecules were unbound, indicating the same affinity of NusA-NTD for β and ββ'. Around 90% of NusA-NTD molecules were free upon addition of β'. Samples containing β' were turbid, suggesting the presence of oligomers with a higher molecular mass, i.e. the fraction of unbound NusA-NTD might be even higher than the estimated value. A small effect of the β' subunit on NusA-NTD binding, however, cannot be excluded. As no interaction was observed between NusA-NTD and the α or the ω subunit (Supplementary Fig. 4), these results agree with previous findings that NusA-NTD interacts with the β flap region (Fig. 3)[35,36].

We probed NusA-AR2:RNAP interaction with the same approach. The signal intensity of ^{15}N-NusA-AR2 was reduced to background levels in the presence of RNAP^{native} (Fig. 5a). The two dimensional (2D)

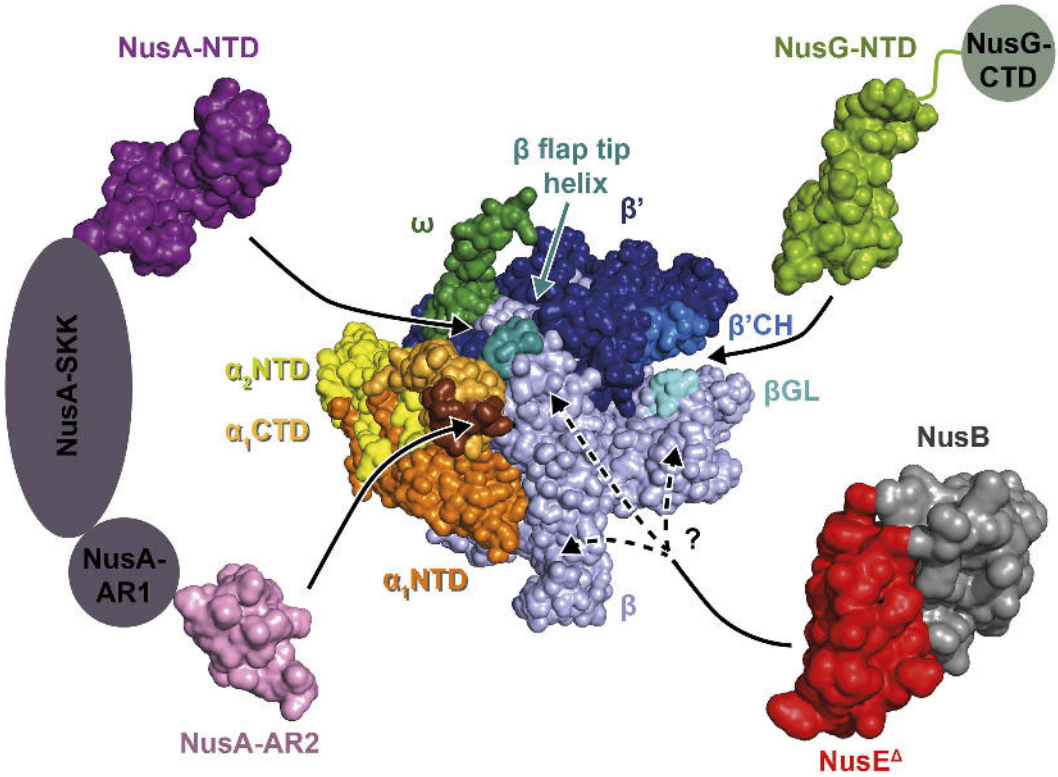


Figure 3. Nus factors binding sites on RNAP. RNAP is shown in surface representation with the NTD and CTD of α subunit 1 in bright and pale orange, respectively, the NTD of α subunit 2 in yellow, the β subunit in pale blue, the β' subunit in dark blue, the ω subunit in dark green. Nus factor binding sites are highlighted (NusA-AR2 binding site on α_1CTD, brown; βGL, cyan; β flap tip helix, turquoise; β'CH, bright blue). Nus factors are displayed in surface representation with linker regions or domains not studied in this work being drawn schematically (NusG, bright green; NusA, purple; NusE$^{\Delta}$, red, NusB, grey). Black arrows indicate the binding site of each Nus factor or domain. NusE$^{\Delta}$ interacts with the β subunit, but the exact binding site has no been identified yet. Protein Data Bank (PDB) codes: RNAP, 4KMU; NusA-NTD, 2KWP; NusA-AR2, 1WCN; NusB:NusE$^{\Delta}$, 3D3B; NusG-NTD, 2K06.

[^{1}H,^{15}N]-HSQC spectrum of ^{15}N-NusA-AR2 changed dramatically when isolated α was added (Fig. 5b), which verifies this interaction. ^{15}N-NusA-AR2 resonances corresponding to amino acid residues known to be located in the αCTD binding surface disappeared[5]. The signal intensity was only slightly diminished in the presence of β, and the spectrum of ^{15}N-NusA-AR2 was completely unaltered upon addition of β' or ω (Fig. 5c,d and Supplementary Fig. 5). Hence, we conclude that NusA-AR2 binds specifically to the α subunit (Fig. 3). Weak binding of NusA-AR1 to β was observed, just as for NusA-AR2. These interactions, however, may be unspecific due to the acidity of the AR domains[33].

Together with the CD spectra these interaction studies suggest that all subunits are functional and consequently correctly folded, although we cannot exclude that regions not interacting with NusA or NusG are not fully intact. Conventional NMR techniques thus allow qualitative studies of the interaction of RNAP with various transcription regulators, setting the stage for further biochemical and structural investigations.

Transcription factor NusE attaches to the β subunit. NusE is able to bind directly to RNAP, an interaction that is suggested to be involved in antitermination[21]. Thus, we asked which RNAP subunit was the target of NusE.

As NusE is only poorly soluble and tends to aggregate, we expressed and purified a NusE variant, NusE$^{\Delta}$, in which the ribosome binding loop is replaced by a single Ser, in complex with NusB[37]. RNAP^{native} or the individual RNAP subunits were added to the NusB:^{15}N-NusE$^{\Delta}$ complex. While addition of α and ω had no effect on the 1D [^{1}H,^{15}N]-HSQC spectrum of ^{15}N-NusE$^{\Delta}$ (Supplementary Fig. 6a,b), RNAP^{native} addition led to a loss of signals indicating binding of NusE$^{\Delta}$ to RNAP (Fig. 6a). A similar signal loss was obtained upon addition of β, demonstrating the formation of the NusE$^{\Delta}$:β complex (Fig. 6b). When β' was added to NusB:^{15}N-NusE$^{\Delta}$, the signal intensity was reduced by approximately 50%, an effect we attribute to weak or unspecific binding (Fig. 6c). To exclude the possibility that NusB alone binds to RNAP we performed a titration experiment with ^{15}N-NusB and RNAP^{native} resulting in an unaltered

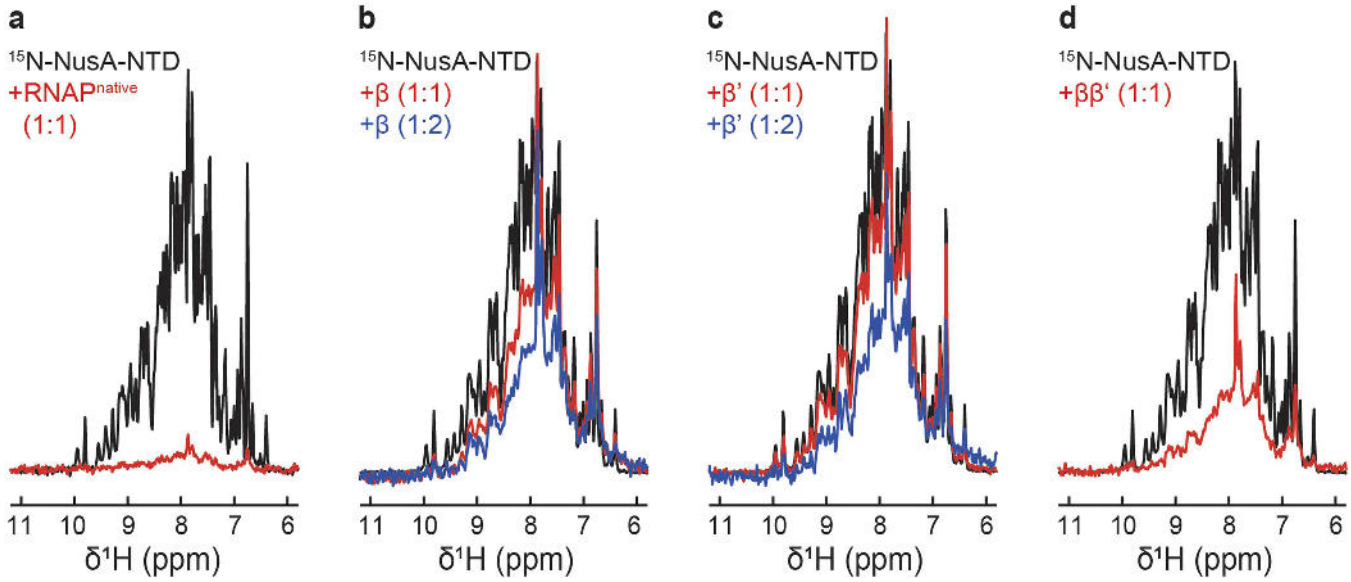


Figure 4. NusA-NTD interaction with RNAP, β, β' and ββ'. 1D [^{1}H,^{15}N]-HSQC spectra of the amide region of 30 μM ^{15}N-NusA-NTD in the absence, black, and in the presence of (**a**) RNAP^{native}, (**b**) β subunit, (**c**) β' subunit, or (**d**) ββ' complex; red, equimolar concentrations; blue, 1:2 molar ratio.

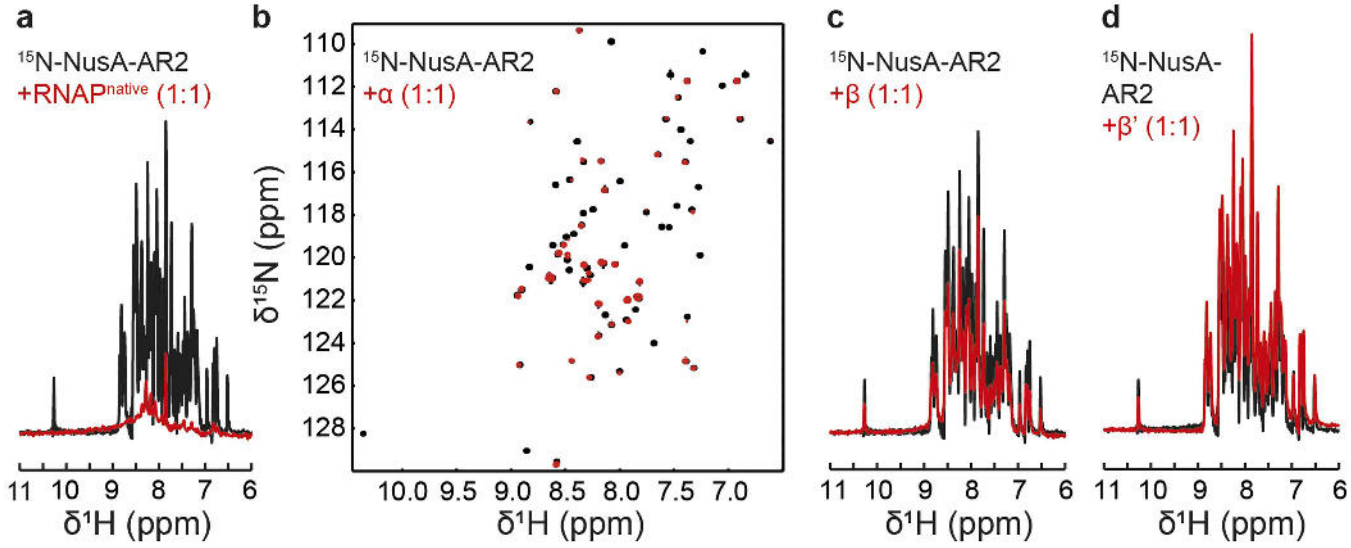


Figure 5. NusA-AR2 interaction with RNAP, α, β, and β'. [^{1}H,^{15}N]-HSQC spectra of the amide region of 30 μM ^{15}N-NusA-AR2 in the absence, black, and in the presence of equimolar concentrations, red, of (**a**) RNAP^{native} (1D spectra), (**b**) α subunit (2D spectra), (**c**) β subunit (1D spectra), or (**d**) β' subunit (1D spectra).

spectrum (Supplementary Fig. 6c). Crosslinking experiments using NusB:NusE$^{\Delta}$ and His$_6$-tagged RNAP^{native} in the presence of paraformaldehyde confirmed the formation of the NusB:NusE$^{\Delta}$:RNAP complex (Supplementary Fig. 6d,e). Addition of RNAP^{native} to ^{15}N-NusB:NusE$^{\Delta}$ led to a dramatic decrease of ^{15}N-NusB signal intensity, indicating that NusB is not released upon binding of NusE$^{\Delta}$ to RNAP (Fig. 6d). Thus, the NusB:NusE$^{\Delta}$ complex directly binds to RNAP *via* NusE$^{\Delta}$ and the β subunit is probably the key target of NusE$^{\Delta}$ (Fig. 3).

Although this might imply that the ribosome could directly interact with RNAP as NusE is part of the 30S subunit, we consider this scenario unlikely as the resulting supramolecular RNAP:ribosome complex would be very rigid and consequently gene expression would probably be impaired. Thus, we propose that the NusE:RNAP interaction might be involved in transcription antitermination as suggested earlier[21].

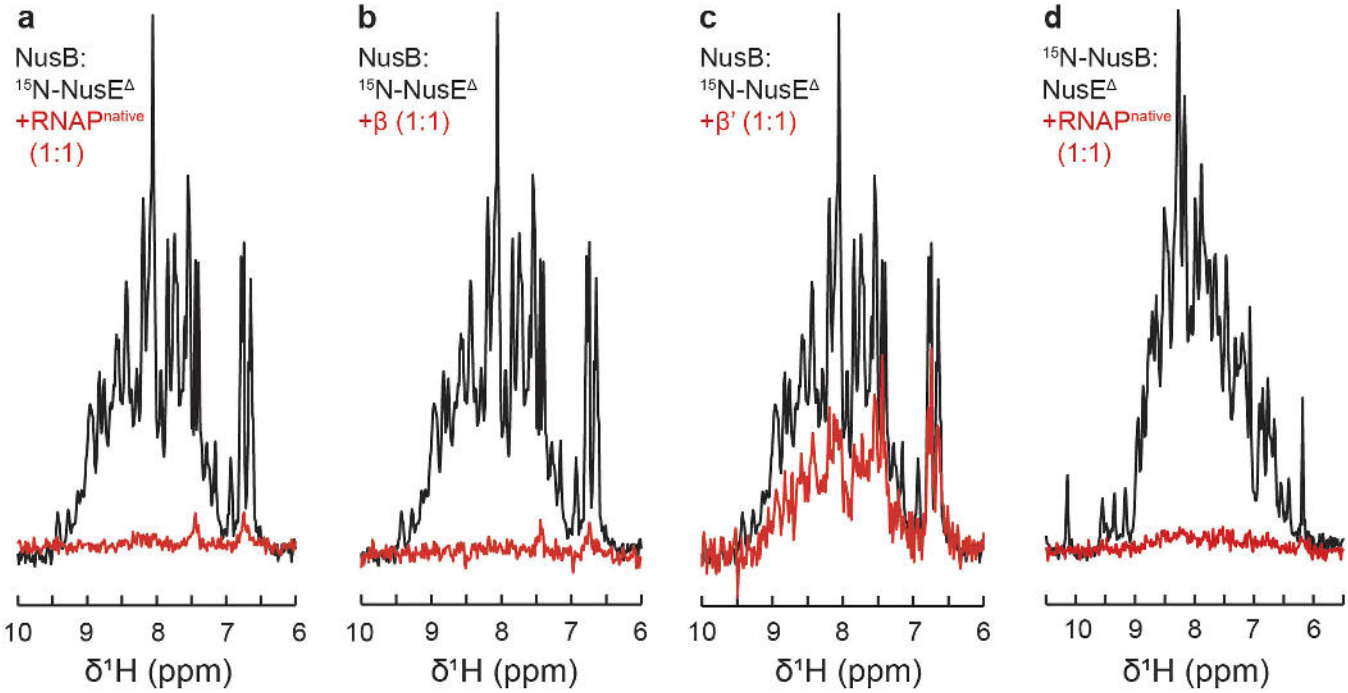


Figure 6. Interaction of NusE$^{\Delta}$ with RNAP, β, and β'. 1D [^{1}H,^{15}N]-HSQC spectra of the amide region of 30 μM NusB:^{15}N-NusE$^{\Delta}$ in the absence, black, and in the presence of equimolar concentrations, red, of (**a**) RNAP^{native}, (**b**) β subunit, or (**c**) β' subunit. (**d**) 1D [^{1}H,^{15}N]-HSQC spectra of the amide region of 30 μM ^{15}N-NusB:NusE$^{\Delta}$ in the absence, black, and in the presence of RNAP^{native} in equimolar concentration, red.

The isolated ω subunit does not interact with the isolated β' subunit. The ω subunit of RNAP was proposed to have an essential function in folding of β' and in preventing it from aggregation as well as in promoting the assembly of $\alpha_2\beta$ with β'ω[7]. The signals of ^{15}N-ω are not diminished significantly upon addition of β', indicating that the two proteins do not interact (Supplementary Fig. 7). Yet, ω co-eluted with the other RNAP subunits in Ni^{2+} affinity chromatography after assembly (Fig. 1b), and ω was present in active RNAP. Thus, we conclude that ω binds only to unfolded or partially folded β'. Together with the analysis of its secondary structure (Fig. 1e and Supplementary Fig. 2) this, in turn, suggests that ω adopts its properly folded state either during RNAP assembly or during folding of β'.

NMR studies of RNAP. The [^{1}H,^{13}C]-TROSY heteronuclear multiple quantum coherence (HMQC) spectrum of deuterated RNAP^{native} with ^{1}H,^{13}C-labeled Ile, Leu, and Val methyl groups shows high signal dispersion, typical for a folded protein (Fig. 7a). However, owing to the size of RNAP (287 Val, 230 Ile, 349 Leu), many signals overlap.

Numerous αCTD signals could be assigned in RNAP^{native} by superposition of a [^{1}H,^{13}C]-HSQC spectrum of ^{13}C,^{15}N-αCTD and the spectrum of RNAP^{native} labeled as above (Fig. 7a), as the αCTD signals in RNAP^{native} are of higher intensity than signals of the rest of the RNAP due to the fact that this domain is flexibly connected to RNAP. A similar approach was used to assign signals in the RNAP^{native} spectrum that belong to the β' subunit (Fig. 7b). In this case, β' was deuterated and contained ^{1}H,^{13}C-labeled methyl groups of Ile, Leu, and Val residues. The signals of the isolated β' subunit are widely dispersed, and several of the RNAP^{native} signals can be assigned clearly to the β' subunit, since the chemical shifts are almost identical in the two spectra.

Addition of unlabeled NusG-NTD to methyl group labeled β' led to a significant decrease of some β' signals (Fig. 7c), indicating that the corresponding residues are affected by NusG-NTD binding. Two Ile and two Leu residues, which give rise to two and four signals in the Ile (^{13}C, 9–16 ppm) and Val/Leu (^{13}C, 17–29 ppm) region, respectively, are positioned directly in the NusG-NTD interaction site of the β'CH (Supplementary Fig. 8), matching the number of significantly affected β' signals. Other Ile, Leu, and Val residues are located in the vicinity of the interaction site and are probably affected by NusG-NTD binding as well (Supplementary Fig. 8). Hence, we conclude that the separately expressed and purified β' subunit is indeed functional in NusG-NTD binding.

In order to reduce the number of signals in the spectrum of methyl group labeled RNAP^{native}, we specifically labeled only the Ile, Val, and Leu methyl groups of the β' subunit with ^{1}H,^{13}C while all other residues of β' as well as the other subunits were deuterated (Fig. 7d). The signals in the resulting spectrum are as well dispersed as the signals of isolated β' (Fig. 7b), but new signals appear. Hence, by comparing the spectrum of methyl group labeled β' in RNAP with the one of methyl group labeled RNAP^{native}

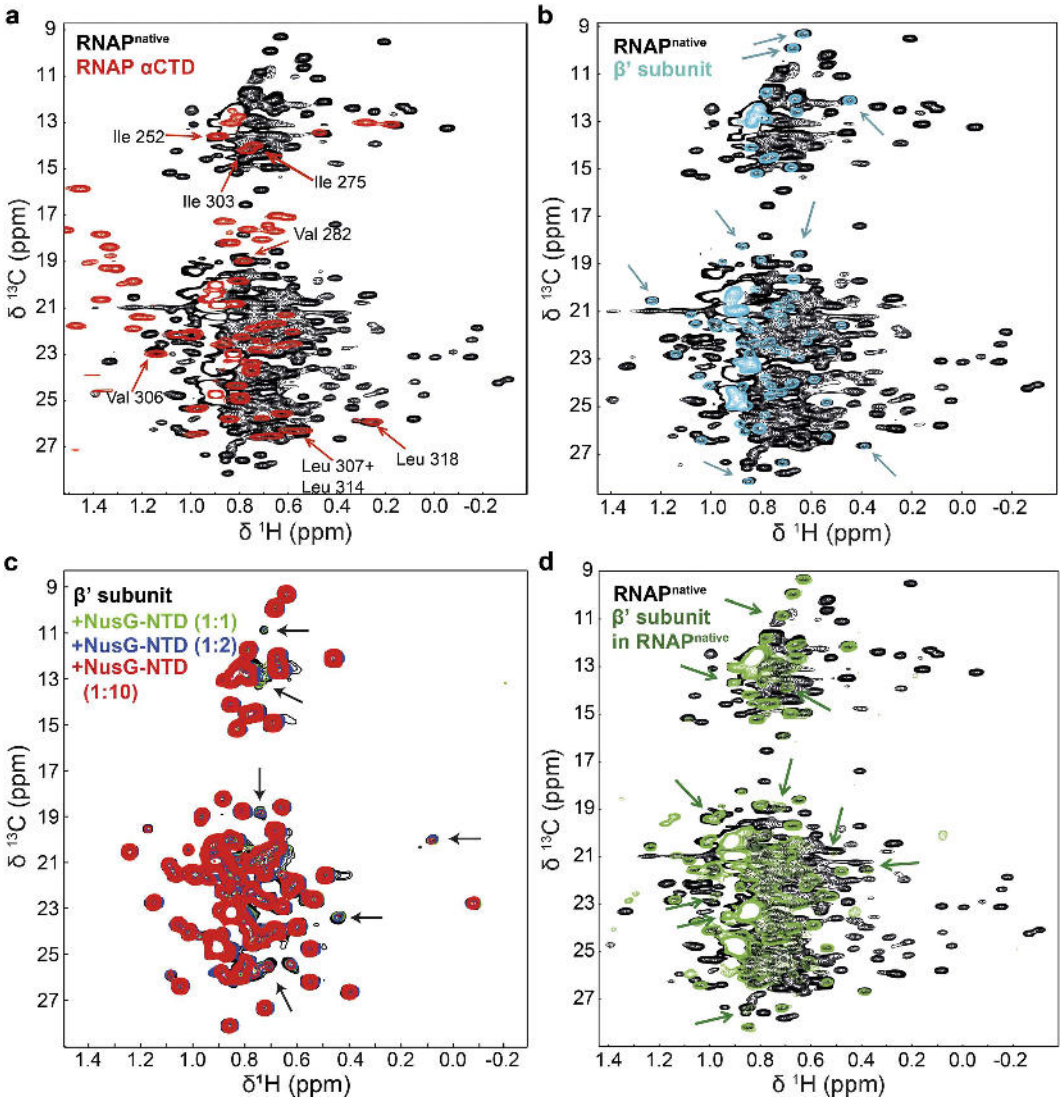


Figure 7. NMR studies of RNAP. C-H correlation spectra of $^{15}N,^{13}C$-RNAP αCTD; methyl group labeled RNAP^{native}; methyl group labeled β'; and methyl group labeled β' in reconstituted RNAP (other subunits deuterated). (**a**) Superposition of a [$^{1}H,^{13}C$]-HMQC spectrum of 30 μM RNAP^{native}, black, and a [$^{1}H,^{13}C$]-HSQC spectrum of 700 μM RNAP αCTD, red. Directly assigned peaks are labeled. (**b**) Superposition of [$^{1}H,^{13}C$]-HMQC spectra of 30 μM RNAP^{native}, black, and 2 μM β', cyan. Example peaks with identical chemical shift in RNAP^{native} and free β' are indicated by blue arrows. (**c**) Superposition of [$^{1}H,^{13}C$]-HMQC spectra of 2 μM β', before, black, and after addition of unlabeled NusG-NTD in a 1:1, 1:2, and 1:10 molar ratio (green, blue, and red, respectively). Arrows indicate signals that decrease significantly upon NusG-NTD addition. (**d**) Superposition of the [$^{1}H,^{13}C$]-HMQC spectra of RNAP^{native}, black, and β' in reconstituted RNAP, green. β' signals identical to signals of RNAP^{native} and those whose positions differ in free β' and β' in reconstituted RNAP are indicated by green arrows.

(Fig. 7d) more signals of RNAP^{native} could be assigned to the β' subunit than using the spectrum of isolated methyl group labeled β'. This is probably due to the fact that here β' was in its physiological context.

Thus, this work demonstrates that even heterooligomeric systems as complex as RNAP can be tackled by NMR spectroscopy, and, moreover, that intra- and interdomain dynamics and the transient interaction with regulatory factors can be studied. In fact, we expect that further refinement of the method we presented here by, e.g., specific labeling of parts of the RNAP subunits will lead to very major contributions to detailed studies of transcription factor:RNAP interactions by solution state NMR spectroscopy.

Methods

Assembly and purification of the RNAP and the ββ′ complex. All RNAP subunit genes were expressed separately (see Supplementary Methods), with the β' subunit being produced as a fusion protein carrying an N-terminal His_6 tag. Cell pellets from equal volumes of cell cultures of the individual

subunits were resuspended in denaturing lysis buffer (50 mM tris(hydroxymethyl)aminomethane (Tris)/HCl, pH 7.5, 500 mM NaCl, 5% (v/v) glycerol, 0.5 mM ethylenediaminetetraacetic acid (EDTA), 10 mM $MgCl_2$, 10 μM $ZnCl_2$, 8 M urea, 1 mM dithiothreitol (DTT)) and combined. Cell lysis was performed with a microfluidizer, and the cell lysate was stirred for 1 h at room temperature. For the assembly of RNAP, the lysate was dialyzed against lysis buffer with decreasing urea concentrations (4 M, 1 M, 0.5 M, 0 M; 2 h each buffer at 4 °C). Finally, the extract was dialyzed overnight against buffer A (50 mM Tris/HCl, pH 7.5, 500 mM NaCl, 5% (v/v) glycerol, 10 mM $MgCl_2$, 10 μM $ZnCl_2$, 10 mM imidazole). The dialysate was incubated for 1 h at 30 °C, centrifuged at 12,000 x *g* and 4 °C for 30 min, and the supernatant was applied to a HisTrap HP column (GE Healthcare, Munich, Germany). After washing with buffer A, elution was performed using a constant gradient with imidazole concentrations increasing up to 1 M in buffer A. RNAP containing fractions were combined and dialyzed against 50 mM Tris/HCl, pH 7.5, 500 mM NaCl, 5% (v/v) glycerol, 0.5 mM EDTA, 10 mM $MgCl_2$, 10 μM $ZnCl_2$, 1 mM DTT at 4 °C overnight. The protein solution was then concentrated by ultrafiltration (molecular weight cut-off (MWCO) = 10 kDa) and applied to a HiLoad 16/600 Superdex 200 pg column (GE Healthcare, Munich, Germany). The fractions of the main peaks from the SEC were concentrated separately by ultrafiltration (MWCO = 10 kDa), frozen in liquid nitrogen and stored at -80 °C.

The assembly and purification of ββ' was performed according to the protocol used for RNAP. However, the incubation step after removing urea was omitted and 37 mg protein were obtained from 1 l cultures.

Protein production and purification of RNAP^{native}. The genes for all subunits were expressed on the same plasmid from one promoter as an operon. Expression and purification are based on a slightly modified published protocol[38]. For the overexpression, the LB/M9 minimal medium[39,40] supplemented with ampicillin (100 μg/ml) was inoculated with a preculture to an OD_{600} of 0.03 and cells were grown at 37 °C. At $OD_{600} \sim 0.2$ the temperature was lowered to 16 °C. After 90 min, overexpression was induced by 0.5 mM IPTG and cells were grown overnight. The first purification step was performed using Ni-NTA Superflow cartridges (QIAGEN, Hilden) on an ÄKTA purifier system.

Isotopic labeling of proteins. ^{15}N- and ^{15}N-, ^{13}C-labeled proteins were obtained by growing *E. coli* in M9 minimal media[39,40] upon respective addition of $(^{15}NH_4)_2SO_4$ (Campro Scientific, Berlin, Germany) and ^{13}C-D-glucose (Spectra Stable Isotopes, Columbia, MD, USA) as the only nitrogen and carbon source. Expression and purification was the same as for proteins produced in LB medium (see Supplementary Methods).

The protocol for deuteration of proteins in which the methyl groups of Ile, Leu and Val residues are $^{1}H,^{13}C$-labeled is based on a published method[26]. First, cells were slowly accustomed to D_2O (Campro Scientific, Berlin, Germany) in precultures (LB, M9 minimal medium in H_2O, M9 with 25% (v/v), 50% (v/v) and 100% (v/v) D_2O consecutively). In the 100% D_2O preculture and the main culture, deuterated glucose (Campro Scientific, Berlin, Germany) was added as the sole carbon source. The time for gene expression was doubled as compared to expression in H_2O. For methyl group labeling, 60 mg/L cell culture 2-keto-3-d_3-4-^{13}C-butyrate (isoleucine; Eurisotop, St. Aubin Cedex, France) and 100 mg/L cell culture 2-keto-3-methyl-d_3-3-d_1-4-^{13}C-butyrate (valine, leucine; Eurisotop, St. Aubin Cedex, France) were added 1 h prior to induction. To produce completely deuterated proteins without ^{13}C or ^{15}N label the final step was omitted.

RNAP activity assay. As RNAP^{native} and the assembled RNAPs do not contain the σ subunit for binding of a promoter region, a nucleic acid scaffold consisting of a template DNA without a promoter, a non-template DNA, and an RNA primer was used for the activity assay. The 24mer template (T24, 5'-GCCGCGCGCTTGCGGTCTGTCCC-3') and 14mer non-template (NT14, 5'-AACGCCAGACAGGG-3') DNA oligos overlap only partially to form a short downstream duplex DNA. The other end of T24 is complementary to the 16mer RNA primer (R16, 5'-GAGUCUGCGGCGCGCG-3') that is labeled with 6-carboxyfluorescein (6-FAM) at the 5'-end for visualization. These oligonucleotides are identical with the ones used to obtain the crystal structure of *Thermus thermophilus* elongation complex PDB code: 2O5I[41]).

The reactions were carried out in 20 mM Tris/HCl, pH 8.0, 40 mM KCl, 0.1 mM EDTA, 0.1 mM DTT. For a 50 μL reaction, 12 pmol of T24 and 10 pmol of R16 were mixed, heated to 75 °C for 5 min, and cooled to RT. 12 pmol of NT14 were added and incubated for 10 min at RT. 20 pmol RNAP were added and again incubated at RT for 10 min. To start the activity assay, 5 mM $MgCl_2$ and 2.5 μM of each NTP were added and incubated at 37 °C for 5 min. ATP and CTP were added for an RNA extension of 3 nt. When GTP was also added, the RNA was extended by 14 nt. The reaction samples were analyzed on a 20% (w/v) polyacrylamide/8.3 M urea gel and fluorescence was visualized by a Stella Imaging System (raytest, Straubenhardt, Germany). To compare the activities of RNAP^{active} and RNAP^{native}, the intensity of the strongest band from extended RNA was divided by the intensity of non-extended RNA primer.

CD measurements. Far-UV CD spectra were recorded on a Jasco J-810 spectropolarimeter (Jasco, Gross-Umstadt, Germany) with protein concentrations between 0.5 and 10 μM in 10 mM potassium phosphate buffer, pH 7.5. Spectra were accumulated ten times at 20 °C with an increment of 0.2 nm.

Measured ellipticity [Θ] was normalized against the protein concentration c in mM, the path length d in cm and the number of amino acids N according to equation (1).

$$[\Theta]_{MRW} = 100 \cdot [\Theta] \cdot /(c \cdot d \cdot N) \quad (1)$$

NMR spectroscopy. NMR measurements were conducted at 25 °C on Bruker *Avance* 600 MHz, 700 MHz, and 800 MHz spectrometers, the latter two equipped with cryogenically cooled probes. The interaction studies of transcription factors with RNAP^{native} and individual subunits were carried out in 25 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), pH 7.5, 50 mM NaCl, 5% (v/v) glycerol, 0.5 mM EDTA, 1 mM DTT. Methyl group and ^{15}N-labeled proteins were in 25 mM HEPES, pH 7.5, 50 mM NaCl, 5% (v/v) glycerol, 0.5 mM EDTA, 10 mM $MgCl_2$, 10 μM $ZnCl_2$, 1 mM DTT while [^{15}N,^{13}C]-αCTD was in 10 mM potassium phosphate, pH 6.4, 50 mM NaCl, 1 mM β-mercaptoethanol. 2D spectra were visualized and analyzed using NMRView[42], 1D spectra by Matlab (The MathWorks, Inc., Version 7.1.0.183). To compare different 1D [^{1}H,^{15}N]-HSQC spectra, the intensity was divided by the number of scans and the protein concentration.

Transverse relaxation rates of amide protons were determined with two-point measurements, using 1D [^{1}H,^{15}N]-HSQC experiments including a spin echo in the first insensitive nuclei enhancement by polarization transfer (INEPT) step[43]. Samples contained either 40 μM ^{15}N-NusA-NTD or 40 μM ^{15}N-NusA-NTD and an equimolar amount of β, β' or ββ' in 25 mM HEPES, pH 7.5, 50 mM NaCl, 5% (v/v) glycerol, 0.5 mM EDTA, 10 mM $MgCl_2$, 10 μM $ZnCl_2$, 1 mM DTT. For the experiment with free NusA-NTD the difference between the two time points for the spin-echo experiments (Δt) was 10 ms, while it was 5 ms for all other measurements. The population-averaged observed R_2 was determined according to equation (2).

$$R_2^{obs} = x_{unbound} \cdot R_2^{NTD} + (1 - x_{unbound}) \cdot R_2^{NTD+partner} \quad (2)$$

R_2^{NTD} is R_2 of free NusA-NTD and $R_2^{NTD+partner}$ is R_2 of the complex of NusA-NTD and β, β' or ββ'. Thus, the fraction of unbound NusA-NTD ($x_{unbound}$) was calculated using equation (3).

$$x_{unbound} = (R_2^{obs} - R_2^{NTD+partner})/(R_2^{NTD} - R_2^{NTD+partner}) \quad (3)$$

R_2^{NTD} corresponds to $R_2^{obs}$of NusA-NTD and was experimentally determined to 50 s^{-1}. $R_2^{NTD+partner}$ was estimated based on the proportionality of R_2 and the molecular mass ($R_2^{NTD+\beta}$: 500 s^{-1}, $R_2^{NTD+\beta'}$: 500 s^{-1}, $R_2^{NTD+\beta\beta'}$: 1000 s^{-1}).

References

1. Werner, F. & Grohmann, D. Evolution of multisubunit RNA polymerases in the three domains of life. *Nat. Rev. Microbiol.* **9,** 85–98 (2011).
2. Burgess, R. R. Separation and characterization of the subunits of ribonucleic acid polymerase. *J. Biol. Chem.* **244,** 6168–6176 (1969).
3. Dove, S. L. & Hochschild, A. Conversion of the ω subunit of *Escherichia coli* RNA polymerase into a transcriptional activator or an activation target. *Genes Dev.* **12,** 745–754 (1998).
4. Jeon, Y. H., Yamazaki, T., Otomo, T., Ishihama, A. & Kyogoku, Y. Flexible linker in the RNA polymerase α subunit facilitates the independent motion of the C-terminal activator contact domain. *J. Mol. Biol.* **267,** 953–962 (1997).
5. Schweimer, K. *et al.* NusA interaction with the α-subunit of *E. coli* RNA polymerase is *via* the UP-element site and releases autoinhibition. *Structure* **19,** 945–954 (2011).
6. Ito, K., Iwakura, Y. & Ishihama, A. Biosynthesis of RNA polymerase in *Escherichia coli*. III. Identification of intermediates in the assembly of RNA polymerase. *J. Mol. Biol.* **96,** 257–271 (1975).
7. Ghosh, P., Ishihama, A. & Chatterji, D. *Escherichia coli* RNA polymerase subunit ω and its N-terminal domain bind full-length β' to facilitate incorporation into the $\alpha_2\beta$ subassembly. *Eur. J. Biochem.* **268,** 4621–4627 (2001).
8. Mustaev, A. *et al.* Modular organization of the catalytic center of RNA polymerase. *Proc. Natl. Acad. Sci. U. S. A.* **94,** 6641–6645 (1997).
9. Daube, S. S. & von Hippel, P. H. Interactions of *Escherichia coli* σ^{70} within the transcription elongation complex. *Proc. Natl. Acad. Sci. U. S. A.* **96,** 8390–8395 (1999).
10. Mukhopadhyay, J. *et al.* Translocation of σ^{70} with RNA polymerase during transcription: fluorescence resonance energy transfer assay for movement relative to DNA. *Cell* **106,** 453–463 (2001).
11. Mooney, R. A., Artsimovitch, I. & Landick, R. Information processing by RNA polymerase: recognition of regulatory signals during RNA chain elongation. *J. Bacteriol.* **180,** 3265–3275 (1998).
12. Artsimovitch, I. & Landick, R. Pausing by bacterial RNA polymerase is mediated by mechanistically distinct classes of signals. *Proc. Natl. Acad. Sci. U. S. A.* **97,** 7090–7095 (2000).
13. Martinez-Rucobo, F. W., Sainsbury, S., Cheung, A. C. & Cramer, P. Architecture of the RNA polymerase-Spt4/5 complex and basis of universal transcription processivity. *EMBO J.* **30,** 1302–1310 (2011).
14. Sevostyanova, A., Belogurov, G. A., Mooney, R. A., Landick, R. & Artsimovitch, I. The β subunit gate loop is required for RNA polymerase modification by RfaH and NusG. *Mol. Cell* **43,** 253–262 (2011).
15. Borukhov, S., Lee, J. & Laptenko, O. Bacterial transcription elongation factors: new insights into molecular mechanism of action. *Mol. Microbiol.* **55,** 1315–1324 (2005).
16. Roberts, J. W., Shankar, S. & Filter, J. J. RNA polymerase elongation factors. *Annu. Rev. Microbiol.* **62,** 211–233 (2008).
17. Santangelo, T. J. & Artsimovitch, I. Termination and antitermination: RNA polymerase runs a stop sign. *Nat. Rev. Microbiol.* **9,** 319–329 (2011).
18. Burmann, B. M. *et al.* A NusE:NusG complex links transcription and translation. *Science* **328,** 501–504 (2010).

19. Mooney, R. A., Schweimer, K., Rösch, P., Gottesman, M. E. & Landick, R. 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 (2009).
20. Friedman, D. I., Schauer, A. T., Baumann, M. R., Baron, L. S. & Adhya, S. L. Evidence that ribosomal protein S10 participates in control of transcription termination. *Proc. Natl. Acad. Sci. U. S. A.* **78,** 1115–1118 (1981).
21. Mason, S. W. & Greenblatt, J. 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 (1991).
22. Sekine, S., Tagami, S. & Yokoyama, S. Structural basis of transcription by bacterial and eukaryotic RNA polymerases. *Curr. Opin. Struct. Biol.* **22,** 110–118 (2012).
23. LeMaster, D. M. & Richards, F. M. NMR sequential assignment of *Escherichia coli* thioredoxin utilizing random fractional deuteriation. *Biochemistry* **27,** 142–150 (1988).
24. Salzmann, M., Pervushin, K., Wider, G., Senn, H. & Wüthrich, K. TROSY in triple-resonance experiments: new perspectives for sequential NMR assignment of large proteins. *Proc. Natl. Acad. Sci. U. S. A.* **95,** 13585–13590 (1998).
25. Tugarinov, V. & Kay, L. E. Quantitative NMR studies of high molecular weight proteins: application to domain orientation and ligand binding in the 723 residue enzyme malate synthase G. *J. Mol. Biol.* **327,** 1121–1133 (2003).
26. Sprangers, R. & Kay, L. E. Quantitative dynamics and binding studies of the 20S proteasome by NMR. *Nature* **445,** 618–622 (2007).
27. Heil, A. & Zillig, W. Reconstitution of bacterial DNA-dependent RNA polymerase from isolated subunits as a tool for the elucidation of the role of the subunits in transcription. *FEBS Lett.* **11,** 165–168 (1970).
28. Borukhov, S. & Goldfarb, A. Recombinant *Escherichia coli* RNA polymerase: purification of individually overexpressed subunits and *in vitro* assembly. *Protein Expr. Purif.* **4,** 503–511 (1993).
29. Tang, H., Severinov, K., Goldfarb, A. & Ebright, R. H. Rapid RNA polymerase genetics: one-day, no-column preparation of reconstituted recombinant *Escherichia coli* RNA polymerase. *Proc. Natl. Acad. Sci. U. S. A.* **92,** 4902–4906 (1995).
30. Gentry, D., Xiao, H., Burgess, R. & Cashel, M. The ω subunit of *Escherichia coli* K-12 RNA polymerase is not required for stringent RNA control *in vivo*. *J. Bacteriol.* **173,** 3901–3903 (1991).
31. Mukherjee, K., Nagai, H., Shimamoto, N. & Chatterji, D. GroEL is involved in activation of *Escherichia coli* RNA polymerase devoid of the ω subunit *in vivo*. *Eur. J. Biochem.* **266,** 228–235 (1999).
32. Katayama, A., Fujita, N. & Ishihama, A. Mapping of subunit-subunit contact surfaces on the β'subunit of *Escherichia coli* RNA polymerase. *J. Biol. Chem.* **275,** 3583–3592 (2000).
33. Eisenmann, A., Schwarz, S., Prasch, S., Schweimer, K. & Rösch, P. The *E. coli* NusA carboxy-terminal domains are structurally similar and show specific RNAP- and λN interaction. *Protein Sci.* **14,** 2018–2029 (2005).
34. Worbs, M., Bourenkov, G. P., Bartunik, H. D., Huber, R. & Wahl, M. C. An extended RNA binding surface through arrayed S1 and KH domains in transcription factor NusA. *Mol. Cell* **7,** 1177–1189 (2001).
35. Toulokhonov, I., Artsimovitch, I. & Landick, R. Allosteric control of RNA polymerase by a site that contacts nascent RNA hairpins. *Science* **292,** 730–733 (2001).
36. Yang, X. *et al.* The structure of bacterial RNA polymerase in complex with the essential transcription elongation factor NusA. *EMBO Rep.* **10,** 997–1002 (2009).
37. Luo, X. *et al.* Structural and functional analysis of the *E. coli* NusB-S10 transcription antitermination complex. *Mol. Cell* **32,** 791–802 (2008).
38. Artsimovitch, I., Svetlov, V., Murakami, K. S. & Landick, R. Co-overexpression of *Escherichia coli* RNA polymerase subunits allows isolation and analysis of mutant enzymes lacking lineage-specific sequence insertions. *J. Biol. Chem.* **278,** 12344–12355 (2003).
39. Meyer, O. & Schlegel, H. G. Biology of aerobic carbon monoxide-oxidizing bacteria. *Annu. Rev. Microbiol.* **37,** 277–310 (1983).
40. Sambrook, J., Fritsch, E. F. & Maniatis, T. in *Molecular Cloning - A Laboratory Manual* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1994).
41. Vassylyev, D. G., Vassylyeva, M. N., Perederina, A., Tahirov, T. H. & Artsimovitch, I. Structural basis for transcription elongation by bacterial RNA polymerase. *Nature* **448,** 157–162 (2007).
42. Johnson, B. A. Using NMRView to visualize and analyze the NMR spectra of macromolecules. *Methods Mol. Biol.* **278,** 313–352 (2004).
43. Iwahara, J., Tang, C. & Marius Clore, G. Practical aspects of ^{1}H transverse paramagnetic relaxation enhancement measurements on macromolecules. *J. Magn. Reson.* **184,** 185–195 (2007).

Acknowledgements

We thank Ramona Heißmann and Ulrike Persau for excellent technical support and Irina Artsimovitch for providing plasmids and for helpful discussions. The work was supported by grant Ro 617/17-1 (to P.R.) from the Deutsche Forschungsgemeinschaft.

Author Contributions

P.R. initiated and supervised the project. B.M.W., J.D., M.S. designed the cloning strategies. B.M.W., J.D., M.S., S.H.K. designed the expression and purification strategies, the CD experiments and the activity assays. J.D. and S.K. designed the crosslinking experiments. J.D., K.S., M.S., S.H.K. designed the NMR experiments. J.D. and M.S. conducted the experiments. J.D., M.S., B.M.W., P.R., S.H.K. wrote the manuscript.

Additional Information

Supplementary information accompanies this paper at http://www.nature.com/srep

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Drögemüller, J. *et al.* Exploring RNA polymerase regulation by NMR spectroscopy. *Sci. Rep.* **5**, 10825; doi: 10.1038/srep10825 (2015).

Supplementary Information

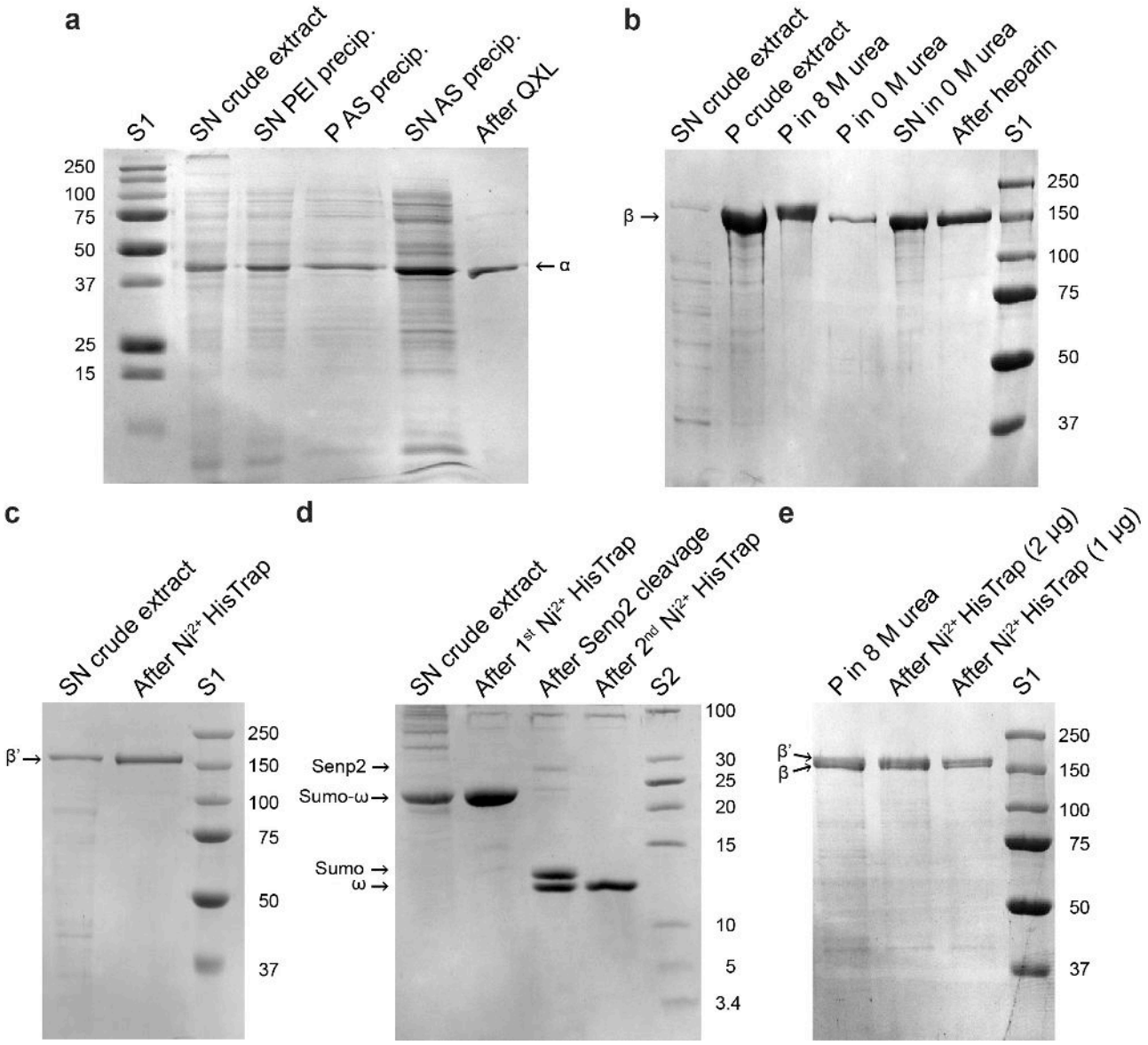
Exploring RNA polymerase regulation by NMR spectroscopy

Johanna Drögemüller[1,†], Martin Strauß[1,†], Kristian Schweimer[1], Birgitta M. Wöhrl[1], Stefan H. Knauer[1,*] & Paul Rösch[1]

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

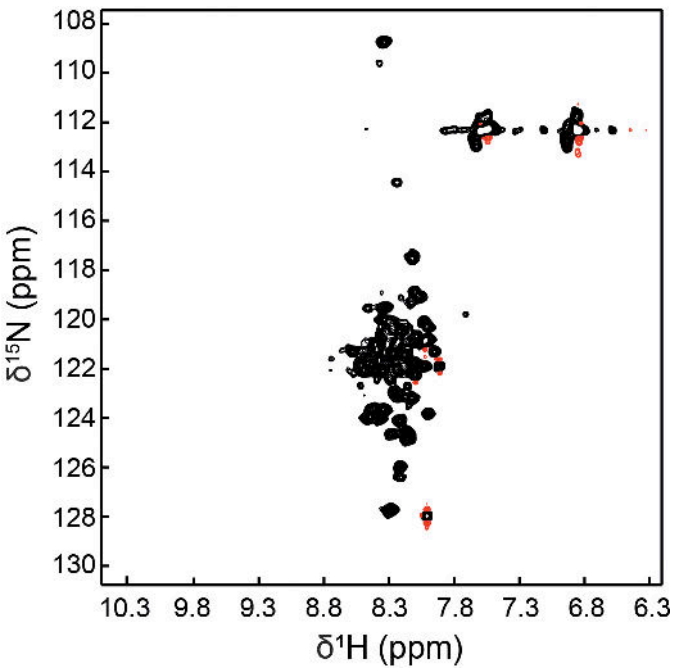
[†] These authors contributed equally to this work

Contents

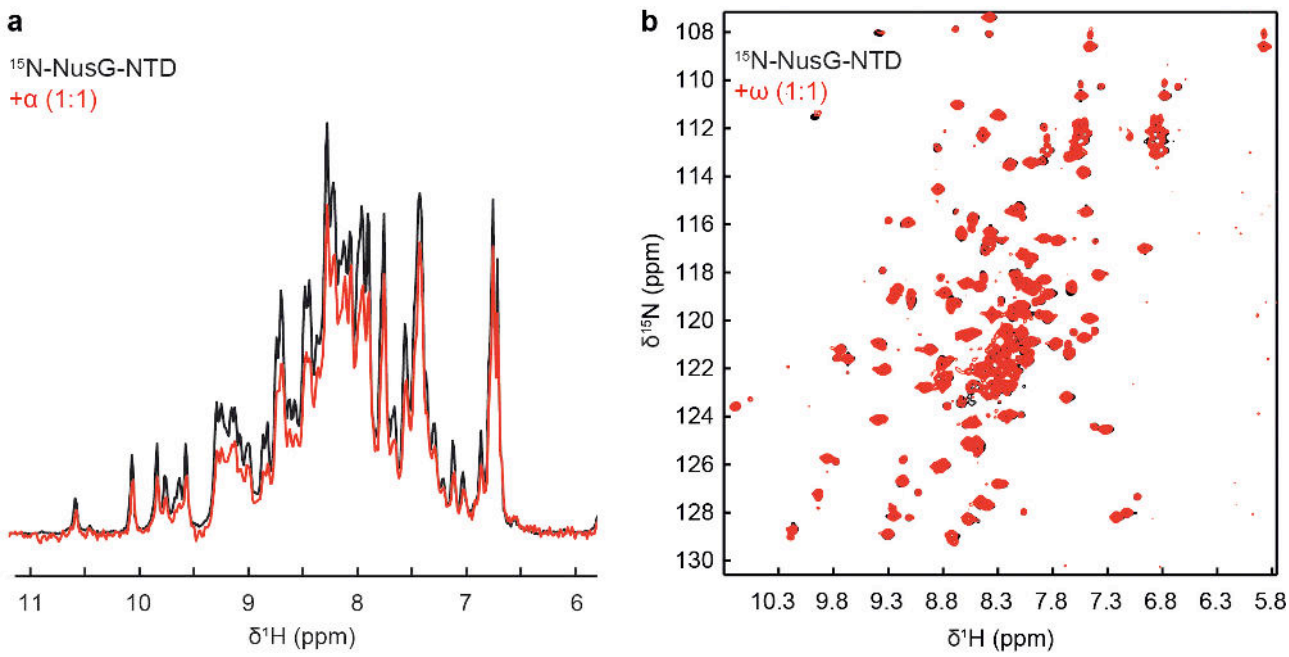


Supplementary Figure 1: Purification of individual RNAP subunits. 2 µg protein were applied to each lane. S1, Precision Plus Protein Standard (BioRad, Munich, Germany); S2, PageRuler Low Range Protein Ladder (Thermo Scientific, Schwerte, Germany); SN, supernatant; P, pellet; PEI, polyethylenimine, AS, ammonium sulfate (**a**) 19 % (w/v) SDS-polyacrylamide gel of aliquots taken from the fractions during α subunit purification after staining with Coomassie Blue. (**b**) 10 % (w/v) SDS-polyacrylamide gel of aliquots taken from the fractions during β subunit purification after staining with Coomassie Blue. (**c**) 10 % (w/v) SDS-polyacrylamide gel of aliquots taken from the fractions during β' subunit purification after staining with Coomassie Blue. (**d**) Schägger-Jagow gel[1] of aliquots taken from the fractions during ω subunit purification after staining with Coomassie

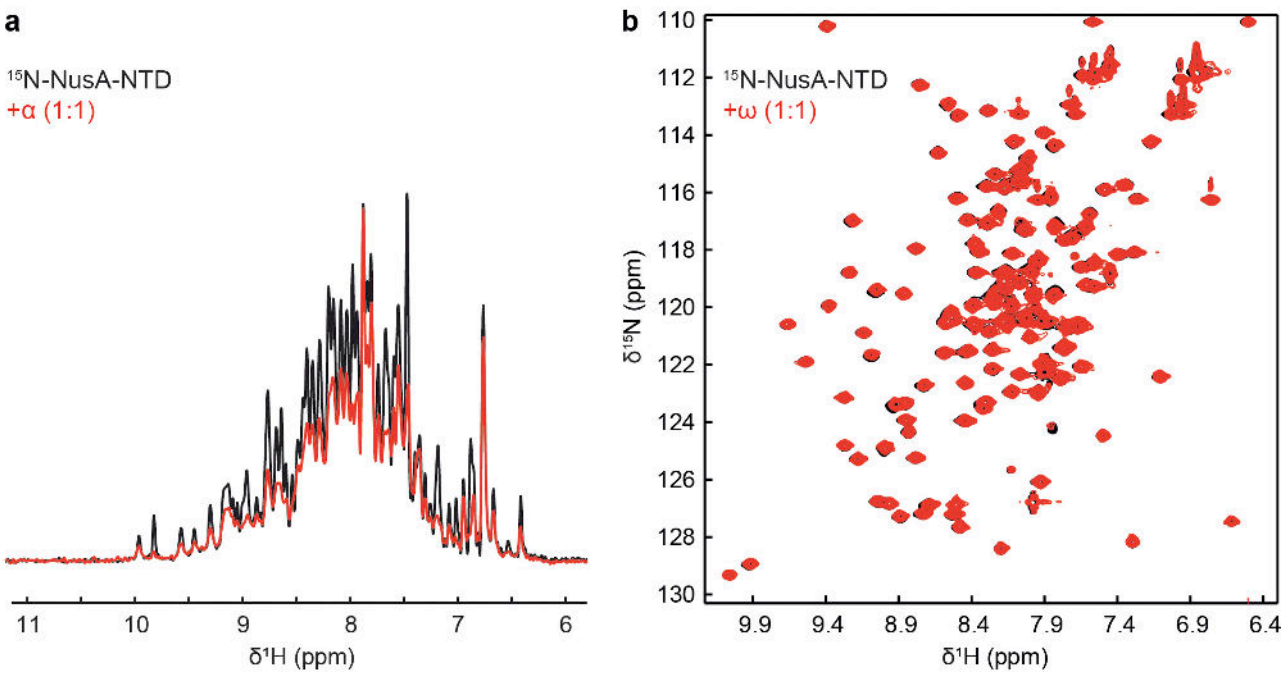
Blue. (**e**) 10 % (w/v) polyacrylamide gel of aliquots taken from the fractions during ββ' complex purification after staining with Coomassie Blue.



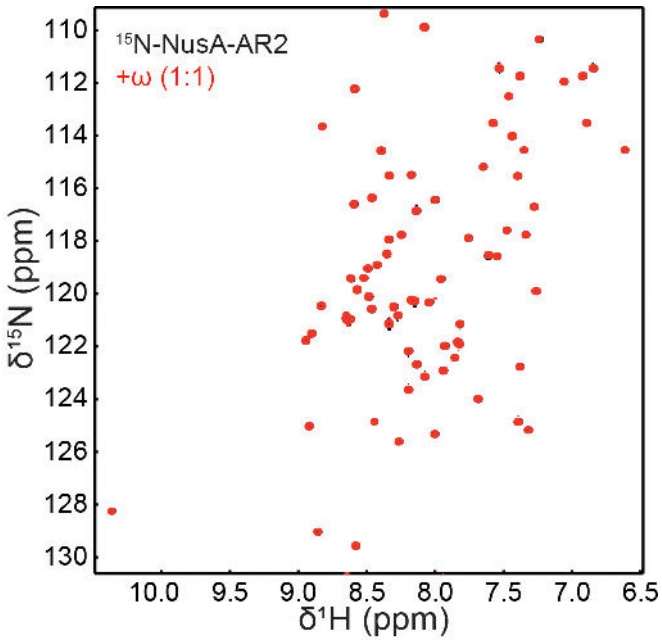
Supplementary Figure 2: [^{1}H,^{15}N]-HSQC spectrum of 300 μM ^{15}N-ω. Positive and negative signals are colored in black and red, respectively.



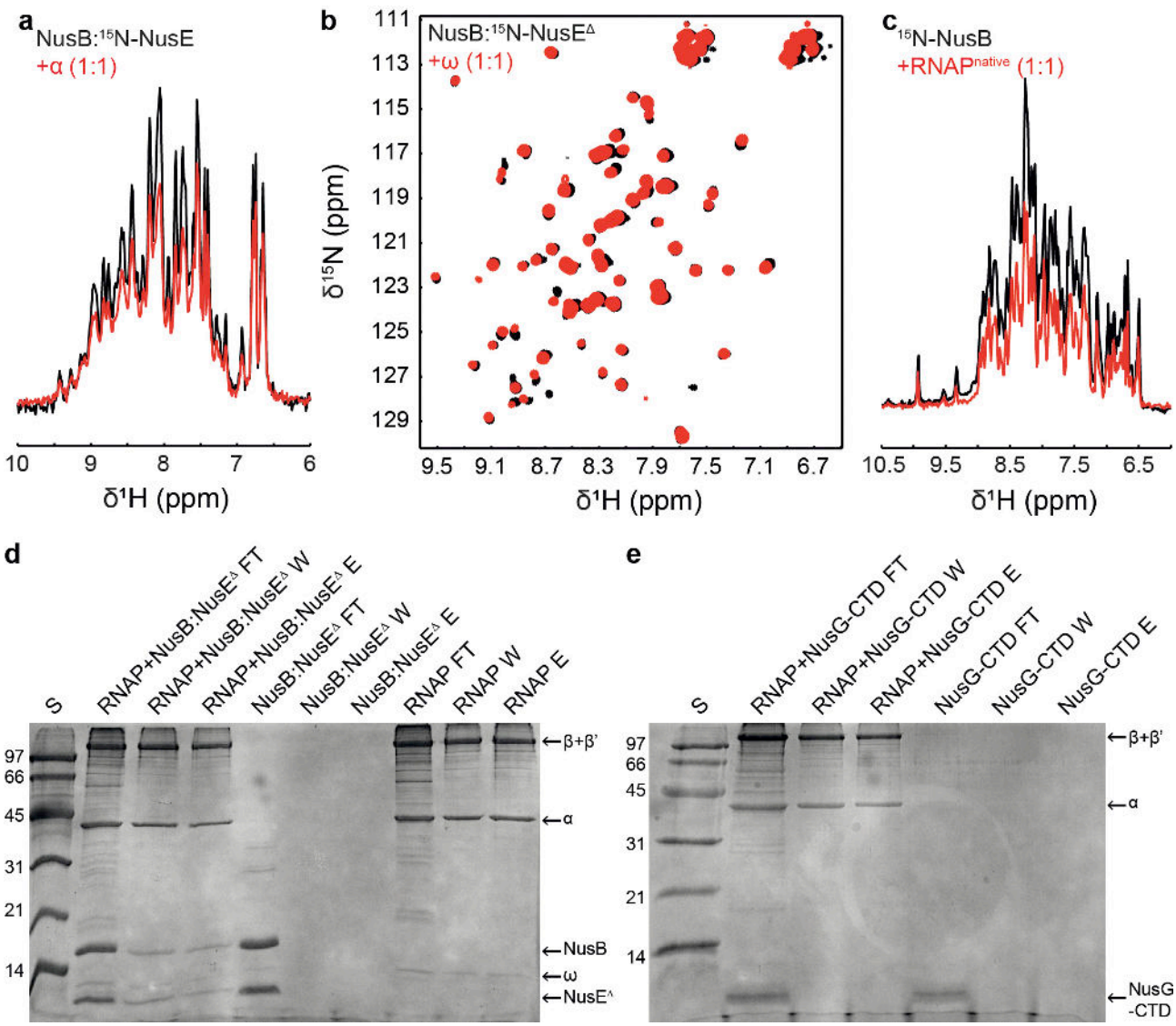
Supplementary Figure 3: NusG-NTD does not interact with isolated α or ω. (**a**) 1D [^{1}H,^{15}N]-HSQC spectra of 30 μM NusG-NTD in the absence, black, and in the presence, red, of an equimolar concentration of α. (**b**) 2D [^{1}H,^{15}N]-HSQC spectra of 100 μM NusG-NTD in the absence, black, and in the presence, red, of an equimolar concentration of ω.



Supplementary Figure 4: NusA-NTD does not interact with isolated α or ω. (**a**) 1D [^{1}H,^{15}N]-HSQC spectra of 30 μM NusA-NTD in the absence, black, and in the presence, red, of an equimolar concentration of α. (**b**) 2D [^{1}H,^{15}N]-HSQC spectra of 100 μM NusA-NTD in the absence, black, and in the presence, red, of an equimolar concentration of ω.

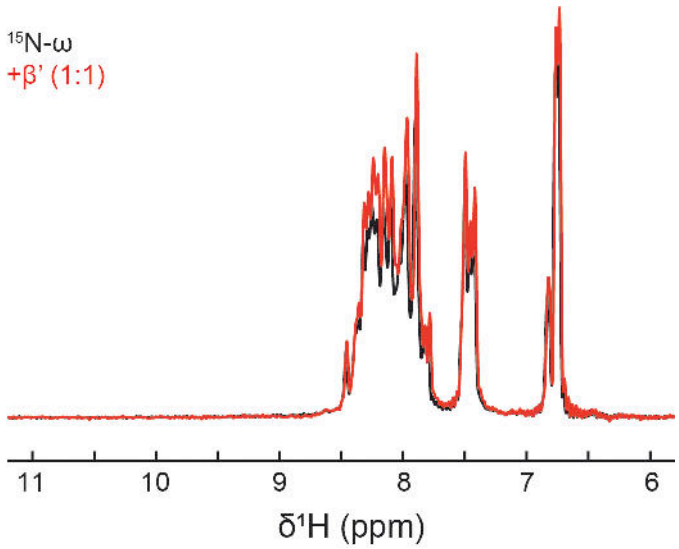


Supplementary Figure 5: NusA-AR2 does not interact with isolated ω. 2D [^{1}H,^{15}N]-HSQC spectra of 30 μM NusA-AR2 in the absence, black, and in the presence, red, of an equimolar concentration of ω.

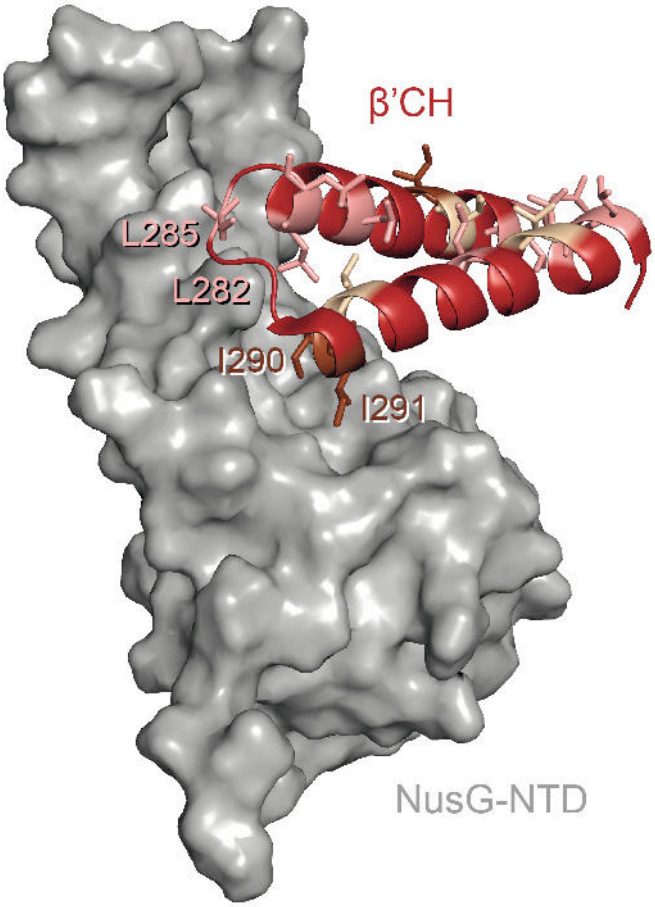


Supplementary Figure 6: Interaction studies of NusB and NusB:NusE$^{\Delta}$ with RNAP, isolated α and ω. (**a**) 1D [^{1}H,^{15}N]-HSQC spectra of 30 μM NusB:^{15}N-NusE$^{\Delta}$ in the absence, black, and in the presence, red, of an equimolar concentration of α. (**b**) 2D [^{1}H,^{15}N]-HSQC spectrum of 100 μM NusB:^{15}N-NusE$^{\Delta}$ in the absence, black, and in the presence, red, of an equimolar concentration of ω. (**c**) 1D [^{1}H,^{15}N]-HSQC spectra of 30 μM ^{15}N-NusB in the absence, black, and in the presence, red, of an equimolar concentration of RNAP^{native}. (**d,e**) Crosslinkling of RNAP and NusB:NusE$^{\Delta}$. 19 % (w/v) SDS-polyacrylamide gel after Ni^{2+} affinity chromatography and staining with Coomassie Blue. Crosslinking of RNAP and NusG-CTD was used as negative control. S: BioRad low range SDS-PAGE Standard (BioRad, Munich, Germany), FT: flow through, W: fraction of the

last washing step with 5 mM imidazole, E: eluate.



Supplementary Figure 7: The ω subunit does not interact with the β' subunit. 1D [^{1}H,^{15}N]-HSQC spectra of the amide region of 30 μM ^{15}N-ω subunit in the absence, black, and in the presence of an equimolar concentration of β', red.



Supplementary Figure 8: Model of NusG-NTD binding to the β'CH. The NusG-NTD: β'CH complex (PDB code: 2K06, NusG-NTD, surface representation; PDB code: 4KMU, β' clamp helices, ribbon representation) was modeled based on the crystal structure of *Pyrococcus furiosus* Spt4/5 binding to the RNAP clamp domain[2]. Ile (brown), Leu (pink) and Val (beige) residues in the β' clamp helices are represented as sticks.

Supplementary Methods

Cloning. Plasmids containing the genes *rpoA*, *rpoB*, *rpoC* and *rpoZ* were kindly provided by Irina Artsimovitch. *rpoB* was cloned from pIA942 into pET29b (Novagen, Madison, WI, USA) *via Bam*HI and *Nde*I. *rpoC* was cloned from pIA661 into pET29b *via Nde*I and *Hind*III restriction sites allowing the expression of *rpoC* with a hexahistidine tag at the C-terminus. For tagless production of *rpoZ* the gene was excised from pIA839 with its ribosome binding site *via Xba*I and *Hind*III and cloned into pET32a (Novagen, Madison, WI, USA). For expression of *rpoZ* with an N-terminal SUMO tag the *rpoZ* gene was cloned into pET28 derivative harboring the small ubiquitin-like modifier (SUMO) 1 gene *via Bam*HI and *Xho*I restriction sites.

The gene for NusA-NTD (1-125) was cloned using the Champion™ pET101 Directional TOPO© Expression Kit (Invitrogen, Carlsbad, CA, USA) with the following primers: Fwd-primer: 5'-CAC CAT GAA CAA AGA AAT TTT GGC-3'; Rev-primer: 5'-AGA ACC ACG CGG AAC CAG CAT CGC ACG TTC GGC TTC ACG-3'. The resulting *E. coli* expression vector pET101_NusA-NTD contains a C-terminal hexa-histidine tag and a thrombin cleavage site between NusA-NTD and the histidine tag.

Gene expression and protein purification. *rpoA* was expressed in *E. coli* BL21 (DE3) (Novagen, Madison, WI, USA) harboring the plasmid pIA287. Cells were grown in M9 minimal medium[3,4] containing 100 µg/ml ampicillin at 37 °C. At an optical density of 600 nm (OD_{600}) of ~ 0.7 expression was induced by 0.1 mM isopropyl-thiogalactoside (IPTG). Cells were harvested after 3 h (9,000 x *g*, 15 min, 4°C), resuspended in 50 mM Tris/HCl (pH 8.0) containing 500 mM NaCl and disrupted by a microfluidizer (Microfluidics, Newton, MA, USA). Nucleic acids were precipitated by addition of 0.6 % (v/v) polyethylenimine and removed by centrifugation (12,000 x *g,* 30 min, 4 °C). Subsequently, an ammonium sulfate precipitation (60 % (w/v)) was performed with the supernatant. After centrifugation (12,000 x *g*, 4 °C, 30 min) the supernatant was dialyzed against 20 mM Tris/HCl (pH 8.0) overnight at 4 °C and applied to a HiTrap QXL column (GE Healthcare,

Munich, Germany). After washing with 20 mM Tris/HCl (pH 8.0) elution was performed using a step gradient with increasing NaCl concentrations (0.25-1 M NaCl in 20 mM Tris/HCl (pH 8.0)). Fractions containing the target protein were combined, dialyzed against the required buffer, concentrated by ultrafiltration (VivaSpin units, molecular weight cut-off (MWCO) = 3.5 kDa, Sartorius Stedim Biotech GmbH, Göttingen, Germany) and stored at -80 °C after freezing with liquid nitrogen. 67 mg protein were obtained from a one liter culture.

rpoB was expressed in *E. coli* BL21 (DE3) harboring the pET29b/*rpoB* plasmid. Cells were grown in M9 minimal medium[3,4] containing 30 µg/ml kanamycin at 37 °C. At an OD_{600} of ~ 0.7 expression was induced by 1 mM IPTG. Cells were harvested 4 h after induction and lysed as described for *rpoA* using a buffer containing 50 mM Tris/HCl (pH 8.0), 500 mM NaCl, 5 % (v/v) glycerol, 1 mM DTT. After centrifugation (30 min, 4 °C, 12,000 x *g*) the pellet was resolved in 1 mM EDTA (pH 8.0), 1 mg/ml deoxycholic acid sodium salt, 20 mM DTT and lysozyme (0.2 mg/ml) and again centrifuged for 30 min at 12,000 x g and 4 °C. The pellet was then washed three times with the same buffer, subsequently three times with 50 mM Tris/HCl (pH 8.0), 50 mM NaCl, 10 mM EDTA, 5 mM DTT and once with H_2O. Finally, the pellet was resuspended in 50 mM Tris/HCl (pH 7.2), 8 M Urea, 500 mM NaCl and stirred for 1 h at room temperature. Urea was removed by dialysis against 50 mM Tris/HCl (pH 7.2), 5 % (v/v) glycerol, 500 mM NaCl, 0.5 mM EDTA, 1 mM DTT at 4 °C for 3 h followed by overnight dialysis using the same buffer without NaCl. The dialysate was centrifuged (30 min, 4 °C, 12,000 x *g*) and the supernatant was applied to a HiTrap Heparin HP column (GE Healthcare, Munich, Germany). After washing with 50 mM Tris/HCl (pH 7.2), 5 % (v/v) glycerol, 0.5 mM EDTA, 1 mM DTT elution was performed using a constant NaCl gradient up to 1 M in 50 mM Tris/HCl (pH 7.2), 5 % (v/v) glycerol, 0.5 mM EDTA, 1 mM DTT. Fractions containing pure β were combined and dialyzed against the required buffer before the protein solution was concentrated by ultrafiltration (MWCO = 10 kDa) and stored at -80 °C after freezing with liquid nitrogen. The yield was 53 mg protein per l culture.

rpoC was expressed in *E. coli* Rosetta (DE3) pLysS (Novagen, Madison, WI, USA). The recombinant protein harbored a seven amino acid linker followed by a hexahistidine tag (His_6) at the C-terminus. Cells were grown in M9 minimal medium[3,4] containing 30 µg/ml kanamycin and 34 µg/ml chloramphenicol at 37 °C. When an OD_{600} of ~ 0.5 was reached the temperature was lowered to 16 °C and gene expression was induced with 1 mM IPTG at an OD_{600} of 0.6-0.8. Cells were harvested 6 h after induction, resuspended and lysed as described above using buffer A (50 mM Tris/HCl (pH 7.5), 500 mM NaCl, 5 % (v/v) glycerol, 10 mM $MgCl_2$, 10 µM $ZnCl_2$, 10 mM imidazole). After centrifugation (30 min, 12,000 x *g*, 4 °C) the supernatant was applied to a HisTrap HP column (GE Healthcare, Munich, Germany). After washing with buffer A, elution was carried out using a step gradient with increasing imidazole concentrations (10-500 mM in buffer A). Fractions containing β' were combined. Following dialysis against the required buffer the protein solution was concentrated by ultrafiltration (MWCO = 10 kDa) and stored at -80 °C after shock freezing in liquid nitrogen. One liter culture yielded 15 mg protein.

The ω subunit with N-terminal His_6-SUMO tag was produced in *E. coli* Rosetta (DE3) pLysS harboring pET28M-SUMO1/*rpoZ.* Cells were grown in M9 minimal medium[3,4] in the presence of 30 µg/ml kanamycin and 34 µg/ml chloramphenicol at 37 °C until an OD_{600} of 0.4 was reached. The temperature was lowered to 25 °C and at an OD_{600} of 0.6-0.8 expression was induced with 1 mM IPTG. Cells were harvested after 4 h, resuspended and lysed as described above. In this case 25 mM Tris/HCl (pH 7.5), 300 mM NaCl, 10 mM imidazole was used for resuspension. After centrifugation (12,000 x *g*, 30 min, 4 °C), the supernatant was applied to a HisTrap HP column. After washing with 25 mM Tris/HCl (pH 7.5), 300 mM NaCl, 10 mM imidazole, elution was performed using a step gradient with increasing imidazole concentrations (10-500 mM in resuspension buffer). Fractions containing His_6-SUMO-ω were combined and cleaved during dialysis overnight against 25 mM Tris/HCl (pH 7.5), 300 mM NaCl by Senp2, a protease that cleaves directly after SUMO protein. The protein solution was reapplied to the HisTrap HP column.

Pure ω was found in the flow through, dialyzed against the required buffer, concentrated by ultrafiltration (MWCO = 3 kDa) and stored at -80 °C after freezing with liquid nitrogen with a yield of 3 mg protein per liter culture.

Tagless ω was used for *in vitro* assembly of RNAP and produced in *E. coli* Rosetta (DE3) pLysS containing pET32a/*rpoZ*. Cells were grown in M9 minimal medium[3,4] containing 100 µg/ml ampicillin and 34 µg/ml chloramphenicol at 37 °C. After induction with 0.1 mM IPTG at an OD_{600} of 0.6-0.8 cells were grown for another 3 h before harvesting (9,000 x *g*, 15 min).

NusA-NTD contained amino acids 1-125 and was produced in *E. coli* BL21 Star (DE3) (Invitrogen, Darmstadt, Germany) harboring pET101_NusA-NTD. Cells were grown at 37 °C in LB medium containing ampicillin (100 µg/ml) until an OD_{600} of 0.6 was reached. Then the temperature was lowered to 20 °C. After 30 min overexpression was induced by 1 mM IPTG. Cells were harvested after overnight growth, resuspended and lysed as described for *rpoA* using a buffer containing 20 mM Tris/HCl (pH 7.5), 150 mM NaCl, 10 % (v/v) glycerol, 20 mM imidazole. After centrifugation at 12,000 x *g* and 4 °C for 30 min, the supernatant was applied to a Ni^{2+}-NTA HiTrap column (GE Healthcare, Munich, Germany). After washing with 20 mM Tris/HCl (pH 7.5), 150 mM NaCl, 10 % (v/v) glycerol, 20 mM imidazole elution was performed *via* a step gradient with increasing imidazole concentrations (20 mM – 1 M imidazole in 20 mM Tris/HCl (pH 7.5), 150 mM NaCl, 10 % (v/v) glycerol, 20 mM imidazole). The fractions containing the NusA-NTD-His_6 fusion protein were combined and the protein was cleaved by thrombin (Novagen, Madison, WI, USA), during dialysis against 20 mM Tris/HCl (pH 7.5) at room temperature overnight. The protein solution was applied to a HiTrap QXL column which was subsequently washed with 20 mM Tris/HCl (pH 7.5) before elution was carried out *via* a step gradient with increasing NaCl concentrations (0 M-1 M NaCl in 20 mM Tris/HCl (pH 7.5)). The fractions containing NusA-NTD were combined and dialyzed against the required buffer. Finally, the protein solution was

concentrated by ultrafiltration (MWCO = 3 kDa), frozen in liquid nitrogen and stored at -80 °C.

The gene of the SUMO protease SENP2 was expressed in *E. coli* Rosetta (DE3) (Novagen, Madison, WI, USA) harboring the plasmid pET28b-senp2. Cells were grown in LB medium containing 30 µg/ml kanamycin and 34 µg/ml chloramphenicol at 37 °C. At OD_{600} ~ 0.7 expression was induced by 1 mM IPTG. Cells were harvested after 4 h (9,000 x *g*, 15 min, 4°C), resuspended in 40 mM Tris/HCl (pH 7.5) containing 500 mM NaCl, 10 mM imidazole and 5 mM DTT and disrupted by a microfluidizer. The supernatant was applied to a HisTrap HP column. Elution was performed using a step gradient with increasing imidazole concentrations (10-500 mM in resuspension buffer). The fractions containing SENP2 were combined, dialyzed against 5 mM Tris/HCl (pH 7.5), 250 mM NaCl, 10 mM DTT, 0.1 mM EDTA and concentrated by ultrafiltration (MWCO = 10 kDa). Finally the glycerol concentration was adjusted to 20 %, aliquots were frozen in liquid nitrogen and stored at -80 °C.

The production and purification of NusB:NusE$^{\Delta}$, NusB, RNAP α-CTD, NusG-NTD, NusG-CTD and NusA-AR2 were carried out as described previously (Refs. [5-7] for NusB:NusE$^{\Delta}$ and NusB, Ref. [8] for αCTD, Ref. [9] for NusG-NTD, Ref. [6] for NusG-CTD, Ref. [8] for NusA-AR2).

Formaldehyde crosslink. The crosslinking of RNAP and NusB:NusE$^{\Delta}$ was based on the SPINE method[10]. 7.7 nmol RNAP were mixed with 15.4 nmol NusB:NusE$^{\Delta}$ in 25 mM HEPES (pH 7.5), 100 mM NaCl and a 4 % (w/v) paraformaldehyde solution in the same buffer was added to a final concentration of 0.6 % (w/v). For the crosslink, the mixture was incubated at 37 °C for 20 min. 0.7 ml of Ni^{2+} chelating sepharose (50 % (w/v), GE Healthcare, Munich, Germany), equilibrated with 25 mM HEPES (pH 7.5), 100 mM NaCl, were added and incubated for 20 min at room temperature. Afterwards the mixture was transferred to a 2.5 ml gravity flow column and the flow trough was collected. The column was washed ten times with 1 ml of 25 mM HEPES (pH 7.5),

100 mM NaCl and seven times with 1 ml of the same buffer containing 5 mM imidazole. Bound protein was eluted with 25 mM HEPES (pH 7.5), 100 mM NaCl, 500 mM imidazole. The protein contained in 200 µl in the flow through, the last washing step and the eluate was precipitated with 50 µl 50 % (v/v) trichloroacetic acid (TCA) by incubation for 20 min on ice and subsequent centrifugation for 10 min at 15,000 x *g*. The pellet was dissolved in 50 µl 2x Roti (Roth, Karlsruhe, Germany). The crosslink was broken by boiling the solution for 20 min and the samples were analyzed by SDS-PAGE. The isolated RNAP, the isolated NusB:NusE$^{\wedge}$ complex as well as NusG-CTD in the absence and presence of RNAP as negative control were treated accordingly.

Programs. All structures were visualized using PyMOL[11].

Supplementary References

1. Schägger, H. & von Jagow, G. Tricine Sodium Dodecyl-Sulfate Polyacrylamide-Gel Electrophoresis for the Separation of Proteins in the Range from 1-Kda to 100-Kda. *Anal. Biochem.* **166,** 368-379 (1987).
2. Martinez-Rucobo, F. W., Sainsbury, S., Cheung, A. C. & Cramer, P. Architecture of the RNA polymerase-Spt4/5 complex and basis of universal transcription processivity. *EMBO J.* **30,** 1302-1310 (2011).
3. Meyer, O. & Schlegel, H. G. Biology of aerobic carbon monoxide-oxidizing bacteria. *Annu. Rev. Microbiol.* **37,** 277-310 (1983).
4. Sambrook, J., Fritsch, E. F. & Maniatis, T. in *Molecular Cloning - A Laboratory Manual* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1994).
5. Burmann, B. M., Luo, X., Wahl, M. C., Rösch, P. & Gottesman, M. E. Fine tuning of the *E. coli* NusB:NusE complex affinity to *BoxA* RNA is required for processive antitermination. *Nucleic Acids Res.* **38,** 314-326 (2010).
6. Burmann, B. M. *et al*. A NusE:NusG complex links transcription and translation. *Science* **328,** 501-504 (2010).
7. Luo, X. *et al*. Structural and functional analysis of the *E. coli* NusB-S10 transcription antitermination complex. *Mol. Cell* **32,** 791-802 (2008).
8. Schweimer, K. *et al*. NusA interaction with the α-subunit of *E. coli* RNA polymerase is via the UP-element site and releases autoinhibition. *Structure* **19,** 945-954 (2011).
9. Burmann, B. M., Schweimer, K., Scheckenhofer, U. & Rösch, P. Domain interactions of the transcription:translation coupling factor *E.coli* NusG are intermolecular and transient. *Biochem. J.* **435,** 783-789 (2011).
10. Herzberg, C. *et al*. SPINE: a method for the rapid detection and analysis of protein-protein interactions *in vivo*. *Proteomics* **7,** 4032-4035 (2007).

11. Schrödinger L. The PyMOL molecular graphics system, version 1.3. *Schrödinger, LLC, Mannheim, Germany* (2010).

7.4 Einzelarbeit D

Johanna Drögemüller*, Martin Strauß*, Kristian Schweimer, Marcel Jurk, Paul Rösch, Stefan H. Knauer (2015): **Determination of RNA polymerase binding surfaces of transcription factors by NMR spectroscopy**, *Scientific Reports,* November 2015.

* Beide Autoren haben in gleichem Maße zur Arbeit beigetragen.

SCIENTIFIC REPORTS

OPEN

Determination of RNA polymerase binding surfaces of transcription factors by NMR spectroscopy

Received: 15 July 2015
Accepted: 13 October 2015
Published: 12 November 2015

Johanna Drögemüller*, Martin Strauß*,‡, Kristian Schweimer, Marcel Jurk†, Paul Rösch & Stefan H. Knauer

In bacteria, RNA polymerase (RNAP), the central enzyme of transcription, is regulated by N-utilization substance (Nus) transcription factors. Several of these factors interact directly, and only transiently, with RNAP to modulate its function. As details of these interactions are largely unknown, we probed the RNAP binding surfaces of *Escherichia coli* (*E. coli*) Nus factors by nuclear magnetic resonance (NMR) spectroscopy. Perdeuterated factors with [^{1}H,^{13}C]-labeled methyl groups of Val, Leu, and Ile residues were titrated with protonated RNAP. After verification of this approach with the N-terminal domain (NTD) of NusG and RNAP we determined the RNAP binding site of NusE. It overlaps with the NusE interaction surface for the NusG C-terminal domain, indicating that RNAP and NusG compete for NusE and suggesting possible roles for the NusE:RNAP interaction, e.g. in antitermination and direct transcription:translation coupling. We solved the solution structure of NusA-NTD by NMR spectroscopy, identified its RNAP binding site with the same approach we used for NusG-NTD, and here present a detailed model of the NusA-NTD:RNAP:RNA complex.

Transcription of genomic information from DNA to RNA is the initial step in gene expression, with RNA polymerase (RNAP) being the key enzyme of this process in all domains of life[1]. Bacterial core RNAP consists of five subunits, 2x α, β, β′, and ω. While the α subunits promote the assembly of the enzyme and are target of many regulatory proteins[2–4], the β and β′ subunits form the active site and catalyze RNA synthesis[5,6]. The ω subunit is supposed to play a structural rather than a functional role. It binds to the N- and C-termini of the β′ subunit to prevent β′ aggregation until the ωβ′ complex is integrated into the RNAP[7]. During initiation of transcription the σ factor binds to core RNAP to form the holo enzyme, and σ is also essential for the recognition and melting of promoter regions (reviewed in[8]). The transcription cycle consists of three major phases: initiation, elongation, and termination. It is highly regulated by a multitude of transcription factors that bind to RNAP modifying its action. Prominent examples are the N utilization substance (Nus) factors that influence especially elongation and termination. Among all transcription factors NusG (Spt5 in archaea and eukaryotes) is unique as it is the only one that is universally conserved[9]. *Escherichia coli* (*E. coli*) NusG is a two-domain protein, with an N-terminal domain (NTD) and a C-terminal domain (CTD) connected *via* a flexible linker[10]. During elongation NusG-NTD binds to RNAP, enhancing the elongation rate and suppressing pauses[10,11]. To fulfill this function NusG-NTD contacts the β′ clamp helices (β′CH) and the β gate loop (βGL), closing the active site cleft so that the nucleic acids are locked and the transcription elongation complex (TEC) is stabilized (Fig. 1)[12,13]. Although NusG/Spt5-NTDs highly likely have the same function in all domains of life, NusG/Spt5-CTDs are targets of various interaction partners and thus serve as recruitment platform for further accessory factors. In *E. coli*, NusG-CTD binds to the termination factor Rho,

Lehrstuhl Biopolymere und Forschungszentrum für Bio-Makromoleküle, Universität Bayreuth, Universitätsstraße 30, 95447 Bayreuth, Germany. †Present address: Max Planck Institute for Molecular Genetics, Ihnestr. 63-73, 14195 Berlin, Germany. ‡Present address: Columbia University Medical Center, New York, NY 10032, USA. *These authors contributed equally to this work. Correspondence and requests for materials should be addressed to S.H.K. (email: stefan.knauer@uni-bayreuth.de)

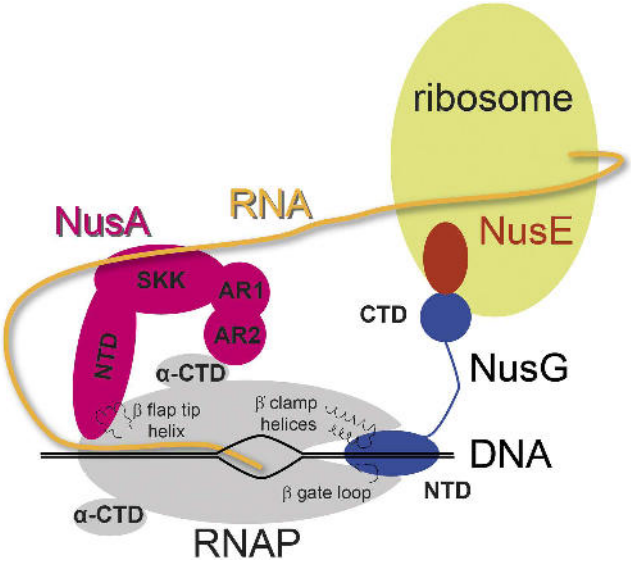


Figure 1. Schematic representation of transcription:translation coupling. NusA, pink, NusE, red; NusG, blue; RNAP, grey; ribosome, light green; DNA, black; RNA, yellow. In RNAP selected structural elements involved in Nus factor binding are indicated.

promoting Rho-dependent termination[14,15]. Additionally, *E. coli* NusG-CTD interacts with ribosomal protein S10 to couple transcription and translation (Fig. 1)[14]. S10 is identical to transcription factor NusE that forms a complex with NusB and as such is involved in antitermination[16]. In the multiprotein antitermination complex RNAP is modified to be able to read through termination signals, a process that is essential for efficient transcription of ribosomal RNA operons[17] or the DNA of lambdoid phages[18]. The NusE:NusB complex formed during antitermination binds to the single stranded, highly conserved *BoxA* RNA sequence[19] and is anchored to RNAP *via* NusE:NusG-CTD interaction[14]. However, NusE is also able to bind directly to RNAP where it remains during elongation[16,20]. This interaction may be involved in antitermination, and the binding site on RNAP is suggested to be located in the β subunit[20].

NusA is a multidomain protein consisting of an NTD, an S1, and two K-homology RNA binding domains, KH1 and KH2, the latter three forming the SKK domain. In *E. coli* and several other proteobacteria the NusA C-terminus comprises two acidic repeat domains, AR1 and AR2[21,22]. With its multitude of interaction partners, NusA is able to accomplish various functions. It modulates Rho-dependent and intrinsic termination, it either prolongs pauses or introduces new ones, and it is part of the antitermination complex (reviewed in[23,24]). NusA interacts directly with RNAP *via* NusA-NTD and NusA-AR2 (Fig. 1)[25]. While a high resolution solution nuclear magnetic resonance (NMR) structure is available for the complex of NusA-AR2 and the CTD of the RNAP α subunit (α-CTD)[4], the RNAP interaction surface of NusA-NTD is not experimentally defined in atomic detail. A low resolution electron microscopy structure of the *Bacillus subtilis* (*B. subtilis*) NusA-NTD:RNAP complex as well as initial binding models are available and all studies suggest that NusA-NTD binds to the flap region of the β subunit at the RNA exit channel[26–28]. However, the exact RNAP binding surface on NusA-NTD remains to be determined.

Knowledge of the RNAP interaction surfaces of transcription factors is crucial for the complete understanding of RNAP regulation. Owed to the molecular mass of RNAP (*E. coli* RNAP ~390 kDa), the main techniques to study RNAP:transcription factor complexes structurally in atomic detail are X-ray crystallography and electron microscopy. However, RNAP regulation heavily depends on transient interactions and dynamics, i.e. information not easily accessible by these techniques. Thus, we chose to study *E. coli* RNAP:Nus factor interaction by NMR spectroscopy to identify the RNAP binding surface of these transcription factors. Our approach is based on observations that even in systems >100 kDa methyl groups are excellent NMR probes as they are still mobile enough to produce highly resolved spectra with good signal intensities owed to their fast motions around the methyl axis[29].

Results and Discussion

RNAP interface of NusG-NTD. To identify the RNAP binding surface of transcription factors the methyl groups of Ile (δ1), Leu (δ1 or δ2), and Val (γ1 or γ2) residues of the respective, deuterated factor were labeled with [^{1}H,^{13}C] ([I,L,V]-labeled transcription factor; for clarity, all protein names without prefix refer to *E. coli* proteins). The titration of this [I,L,V]-labeled regulator with protonated RNAP was observed by two-dimensional (2D) [^{1}H,^{13}C]-methyl transverse relaxation optimized spectroscopy (TROSY). As a test case for the applicability of this method, we asked whether we were able to confirm

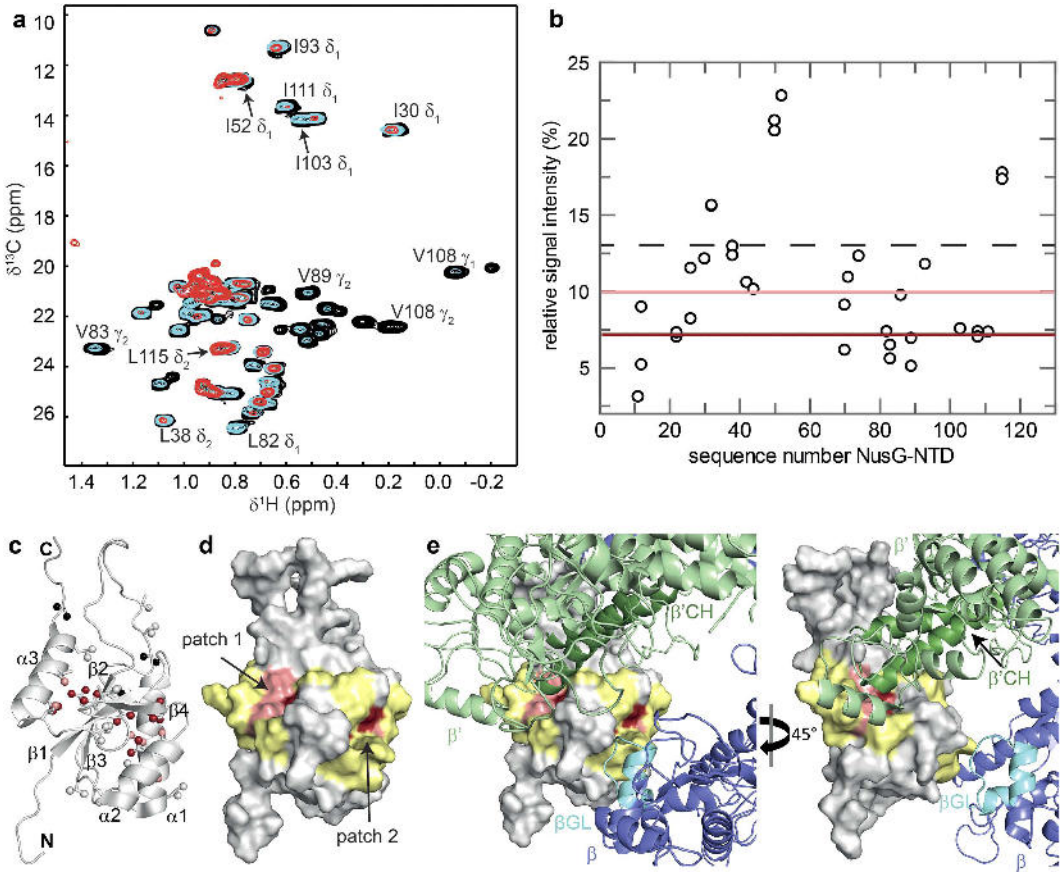


Figure 2. RNAP binding site of NusG-NTD. (**a**) Titration of [I,L,V]-NusG-NTD with protonated RNAP. Methyl-TROSY spectra of [I,L,V]-NusG-NTD in the absence, black, and in the presence of RNAP (1:1 molar ratio, cyan; 1:2 molar ratio, red). Selected signals are labeled. (**b**) Relative signal intensity of [I,L,V]-NusG-NTD after addition of RNAP in equimolar concentration *vs.* residue number of NusG-NTD. The dashed black line indicates the average relative signal intensity. Dark red and light red lines indicate the thresholds for strongly affected (55% of the average relative intensity) and slightly affected (75% of the average relative intensity) methyl groups, respectively. (**c**) Mapping of affected methyl groups onto the NusG-NTD structure (Protein Data Bank (PDB) ID: 2K06, cartoon representation, grey). Ile, Leu, and Val residues are in stick representation with the carbon atoms of their methyl groups as spheres. Strongly affected methyl groups, dark red; slightly affected methyl groups, light red; unaffected methyl groups, grey; unassigned methyl groups, black. Secondary structure elements and termini are labeled. (**d**) Mapping of affected residues onto the NusG-NTD structure (surface representation). For graphical illustration of the interaction site the complete amino acid was colored as affected in lieu of the methyl group. Colors are as in (**c**). Two amino acids on either side of affected Ile/Leu/Val residues are highlighted in yellow unless they were unaffected Ile/Leu/Val residues. (**e**) Model of NusG-NTD as in (**d**) bound to *E. coli* RNAP (PDB ID: 4KMU). The model is based on the structure of the *Pyrococcus furiosus* (*P. furiosus*) Spt4/5 complex bound to the RNAP clamp domain (PDB ID: 3QQC). NusG-NTD was superposed on Spt5 and RNAP β′ subunit on the clamp domain. As NusG-NTD and RNAP were treated as rigid bodies and no further optimization was carried out some minor clashes occur. β subunit, light blue; β′ subunit, light green; β′CH, dark green; βGL, cyan.

the RNAP binding surface of NusG-NTD. This surface is known from a crystallographic study of the archaeal Spt4/5 complex with the β′ clamp domain of RNAP and biochemical experiments on NusG and RfaH, the latter being a paralog of NusG[12,13].

Upon addition of RNAP, the methyl group signals of [I,L,V]-NusG-NTD decreased in intensity, but not uniformly over all signals (Fig. 2a), likely caused by a combination of several effects. First, a general loss of signal intensity is owed to [I,L,V]-NusG-NTD:RNAP complex formation as the molecular mass (MM) of the complex is roughly 30-fold that of [I,L,V]-NusG-NTD ($MM_{NusG\text{-}NTD}$ = 14 kDa, MM_{RNAP} = 389 kDa), resulting in severe line broadening. Second, by binding of [I,L,V]-NusG-NTD to RNAP, the specifically labeled methyl groups of [I,L,V]-NusG-NTD located in the binding interface get into close proximity of the RNAP protons, and the resulting intermolecular dipole-dipole interactions cause an additional contribution to relaxation, so that the signal intensity of methyl groups in the binding surface decreases more strongly than that of methyl groups located elsewhere in [I,L,V]-NusG-NTD.

Finally, signal intensities can be influenced by chemical exchange processes in the intermediate range of the NMR timescale. Quantitative analysis of signal intensities for the 1:1 complex revealed two patches in the protein structure where signal intensities changed noticeably (Fig. 2b,c). Patch 1 comprises residues in helix α3 and strands β1 and β3, while patch 2 is formed by residues located in helices α1 and α2, and these two patches are located at nearly opposite sides of NusG-NTD. No assigned, but unaffected methyl groups were found in either of these patches. This approach provides only information about Ile, Leu, and Val residues, but most likely additional amino acids, especially in the direct vicinity of the affected residues, are involved in the interaction. Thus we graphically extended the representation of patches 1 and 2 by including the two residues preceding and following each affected Ile, Leu, or Val residue, unless they were unaffected Ile, Leu, or Val residues, resulting in two continuous regions (Fig. 2d). In a model of NusG-NTD bound to RNAP based on the crystal structure of the archaeal Spt4/5: β′ clamp domain complex[12], residues of patch 1 are in direct proximity of the β′CH, indicating that we identified correctly the β′CH binding site (Fig. 2e). The NTD of RfaH, an *E. coli* paralog of NusG, not only interacts with the β′CH, but also binds to the βGL *via* His65, Thr66, and Thr67 which form an HTT motif located at the N-terminus of helix α2 (Supplementary Fig. 1)[13]. Although this interaction does not contribute significantly to the overall affinity of RfaH-NTD for RNAP it is essential for the antipausing activity of RfaH[13]. Similarly, structurally homologous residues in NusG-NTD (Ser79-His81) have been proposed to be involved in βGL binding, suggesting that this interaction is a general feature of NusG-like proteins[13]. NusG-NTD patch 2 corresponds to the RfaH region that is in immediate neighborhood of the βGL binding motif suggested for RfaH-NTD (Supplementary Fig. 1)[13]. Due to the absence of Ile, Leu, and Val residues in the NusG-NTD region that is structurally homologous to the HTT motif in RfaH, no direct information about this region is available in our experiments (Supplementary Fig. 1). Thus, we conclude that either the βGL binding surface in NusG-NTD differs slightly from the one in RfaH-NTD or that patch 2 constitutes only part of the βGL interaction surface or that residues of patch 2 are indirectly affected as they are located next to the actual binding site.

The clamp domain undergoes structural rearrangements during the transcription cycle, having closed and open conformations, and NusG-NTD/RfaH-NTD is proposed to lock the clamp in a closed state during elongation by making bridging contacts between the β′CH and the βGL so that the downstream DNA is completely encircled[13,30–33]. Hence, the elongation complex is stabilized and structural rearrangements that occur during pausing are prevented, which, in turn, leads to increased processivity. As we used core RNAP in our experiments the clamp is probably in an open state. Thus our findings indicate that in the absence of nucleic acids NusG-NTD contacts the β′CH and βGL either separately or simultaneously, suggesting that the RNAP claw is in a conformation that allows these contacts or that NusG-NTD induces a closed state.

Overall, the binding surfaces identified here are consistent with the previously published interaction sites of NusG-NTD, demonstrating that the present approach may be used to determine the RNAP binding surfaces of transcription factors in solution in a single experiment using intact RNAP and avoiding molecular alteration of the constituents. However, the limited number of NMR probes and their distribution over the structure restricts the structural resolution of the resulting binding site. Although we are not able to distinguish between methyl groups that are directly involved in the molecular interaction from those that are only indirectly affected, the careful interpretation of the surface representation allows us to identify the interaction surface.

RNAP interface of NusE. Transcription factor NusE/S10 not only interacts with RNAP *via* NusG, but it is also able to bind directly and specifically to the RNAP β subunit during transcription[14,16,20]. The function of this interaction is still unknown. In order to study the molecular details of this interaction we determined the RNAP binding surface of NusE with the same approach as for NusG-NTD. As NusE alone is very unstable and tends to aggregate we used a NusE variant that lacks the ribosome binding loop ($NusE^{\Delta}$) in complex with NusB for our experiments[34]. The presence of NusB does not influence the $NusE^{\Delta}$:RNAP interaction[20]. For the NMR titration, we labeled the methyl groups of Ile, Leu, and Val residues of $NusE^{\Delta}$ in the deuterated NusB:$NusE^{\Delta}$ complex with [^{1}H,^{13}C] ([I,L,V]-$NusE^{\Delta}$).

Upon addition of protonated RNAP, [I,L,V]-$NusE^{\Delta}$ methyl group signals decreased in varying proportion (Fig. 3a,b). All highly and slightly affected methyl groups are located in helices α1 and α2 as well as strands β1 and β4 (Fig. 3c). Inspection of the surface representation and the graphical extension as carried out for NusG-NTD result in a continuous patch (Fig. 3d). As the 7 Ile, 10 Leu, and 7 Val residues of $NusE^{\Delta}$ (86 residues overall) are distributed evenly over the sequence and the structure, our definition of the interaction surface is highly reliable. The RNAP binding site is opposite of the NusB:$NusE^{\Delta}$ interface and the ribosome integration site, i.e. the $NusE^{\Delta}$:RNAP interaction is not only possible within the context of the NusB:$NusE^{\Delta}$ complex, but also when NusE is integrated into the ribosome[35]. NusE could thus simultaneously accommodate the ribosome and the RNAP.

Interestingly, $NusE^{\Delta}$'s binding surface for RNAP strongly overlaps with that for NusG-CTD so that binding of $NusE^{\Delta}$ to RNAP and NusG-CTD should be mutually exclusive (Fig. 3e)[14]. Thus we asked whether NusG-CTD and RNAP compete for binding to NusE. We performed a [^{1}H,^{15}N]-heteronuclear single quantum coherence (HSQC) displacement experiment in which the complex NusB:[^{15}N]-$NusE^{\Delta}$:RNAP was titrated with NusG-CTD (Fig. 4a). In the one-dimensional (1D) [^{1}H,^{15}N]-HSQC spectra signals of [^{15}N]-$NusE^{\Delta}$ strongly decreased upon NusB:[^{15}N]-$NusE^{\Delta}$:RNAP complex formation as the increase

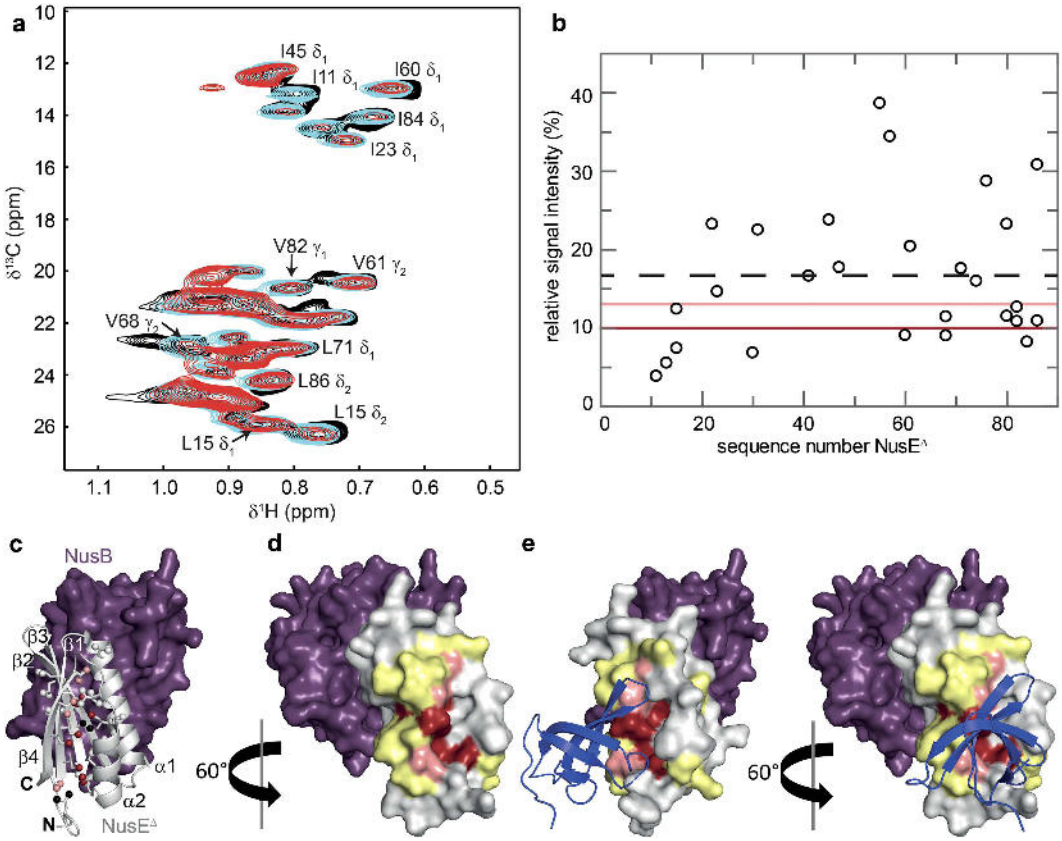


Figure 3. RNAP binding site of NusE$^{\Delta}$. (**a**) Titration of [I,L,V]-NusE$^{\Delta}$ with protonated RNAP (NusE$^{\Delta}$ being in complex with deuterated NusB). Methyl-TROSY spectra in the absence, black, and in the presence of RNAP (1:1 molar ratio, cyan; 1:2 molar ratio, red), with representative signal assignments. (**b**) Relative [I,L,V]-NusE$^{\Delta}$ signal intensity after addition of RNAP in a 1:2 molar ratio *vs.* amino acid sequence positions of NusE$^{\Delta}$. Dashed black line, average relative signal intensity; dark red and light red lines, thresholds for strongly affected (60% of the average relative intensity) and slightly affected (80% of the average relative intensity) methyl groups, respectively. (**c**) Mapping of affected methyl groups onto the NusB:NusE$^{\Delta}$ complex structure (PDB ID: 3D3B; NusB, purple; NusE$^{\Delta}$, light grey). NusB in surface, NusE$^{\Delta}$ in cartoon representation. Ile, Leu, and Val residues in NusE$^{\Delta}$ are represented as sticks with the carbon atoms of their methyl groups as spheres. Strongly affected methyl groups, dark red; slightly affected methyl groups, light red; unaffected methyl groups, grey; unassigned methyl groups, black. Secondary structure elements and termini are labeled. (**d**) Mapping of affected residues onto the NusB:NusE$^{\Delta}$ complex structure (surface representation). Colors are as in (**c**). For graphical illustration of the interaction site the complete amino acid was colored as affected in lieu of the methyl group. Two amino acids on either side of an affected Ile/Leu/Val residue are highlighted in yellow unless they were unaffected Ile/Leu/Val residues. (**e**) Structure of the NusB:NusE$^{\Delta}$:NusG-CTD complex. The NusE$^{\Delta}$:NusG-CTD complex (PDB ID: 2KVQ, NusG-CTD in blue cartoon representation) was superposed on the NusB:NusE$^{\Delta}$ complex from (**d**).

of the molecular mass leads to significant line broadening. Titration with NusG-CTD reversed this effect, demonstrating the displacement of RNAP from NusB:[^{15}N]-NusE$^{\Delta}$. The corresponding 2D [^{1}H,^{15}N]-HSQC spectra show that released NusB:[^{15}N]-NusE$^{\Delta}$ binds to NusG-CTD (Supplementary Fig. 2). Thus, NusG-CTD can abstract NusE$^{\Delta}$ from RNAP. Next, we asked whether in reverse RNAP can displace NusG-CTD from the NusB:NusE$^{\Delta}$:NusG-CTD complex. We titrated NusB:NusE$^{\Delta}$:[^{15}N]-NusG-CTD with RNAP and followed the titration by recording 2D [^{1}H,^{15}N]-HSQC spectra (Fig. 4b,c). Addition of NusB:NusE$^{\Delta}$ to [^{15}N]-NusG-CTD led to changes in the chemical shifts of [^{15}N]-NusG-CTD signals typical for NusB:NusE$^{\Delta}$:[^{15}N]-NusG-CTD complex formation. Those changes were reversed by about 50% when RNAP was added in 3-fold molar excess, as expected on disruption of the NusB:NusE$^{\Delta}$:NusG-CTD complex by NusE:RNAP interaction. Thus, RNAP and NusG-CTD compete for NusE$^{\Delta}$ with similar low micromolar K_D values (NusB:NusE$^{\Delta}$:NusG-CTD: 50 μM)[14].

These competition experiments support the notion of overlapping binding sites of NusE for NusG-CTD and RNAP, and they show that NusG-CTD can interact with NusE in the presence of RNAP. The complexes NusE:RNAP and NusE:NusG:RNAP *via* NusG are thus in a delicate equilibrium that can easily be influenced by other regulators such as transcription factors or certain RNA sequences. Overall, formation

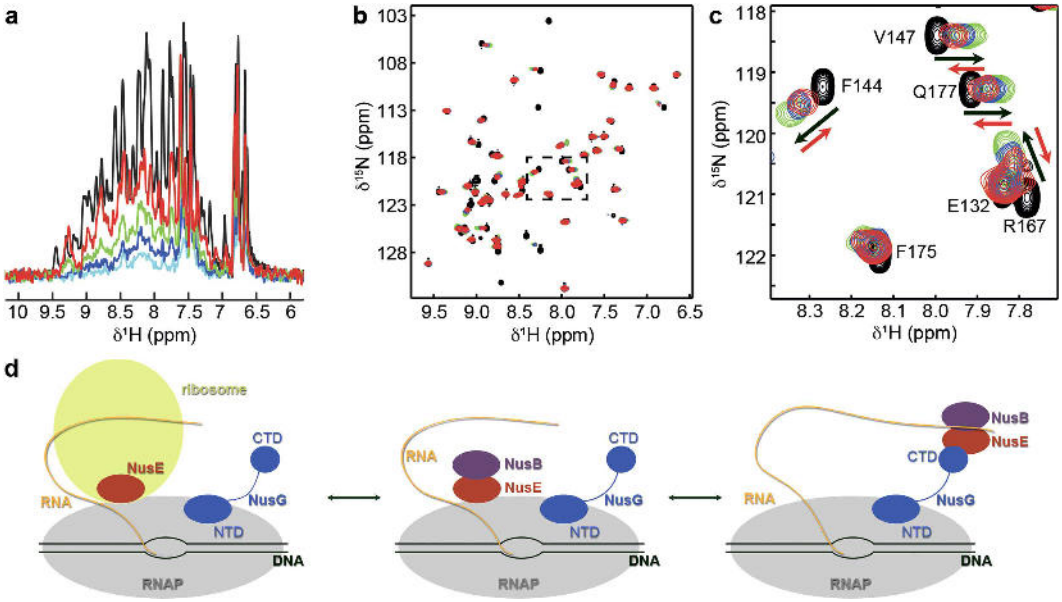


Figure 4. Competition of RNAP and NusG-CTD for NusE binding. (**a**) Displacement of RNAP from NusB:NusE$^{\Delta}$ by NusG-CTD. 1D [^{1}H,^{15}N]-HSQC spectra of free NusB:[^{15}N]-NusE$^{\Delta}$, black, NusB:[^{15}N]-NusE$^{\Delta}$ in the presence of RNAP in equimolar concentration, light blue, and NusB:[^{15}N]-NusE$^{\Delta}$ in the presence of RNAP and NusG-CTD (molar ratio 1:1:1, dark blue; 1:1:3, green; 1:1:10, red). (**b**) Displacement of NusB:NusE$^{\Delta}$ from NusG-CTD by RNAP. 2D [^{1}H,^{15}N]-HSQC spectra of [^{15}N]-NusG-CTD, black, [^{15}N]-NusG-CTD in the presence of NusB:NusE$^{\Delta}$ in equimolar concentration, green, and [^{15}N]-NusG-CTD in the presence of NusB:NusE$^{\Delta}$ and RNAP (molar ratio 1:1:1, blue; 1:1:3, red). (**c**) Detail of the rectangular region in (**b**). Black arrows indicate the chemical shift changes that occur upon addition of NusB:NusE$^{\Delta}$ to [^{15}N]-NusG-CTD, red arrows show the changes upon subsequent addition of RNAP. (**d**) Schematic representation of the potential functions of a direct NusE:RNAP interaction. Color code as in Fig. 1.

of the NusE:RNAP complex might play various roles during transcription (Fig. 4d). It might be involved either in transcription:translation coupling as the ribosome could directly contact RNAP *via* S10, e.g. when the RNA tether is relatively short, or in transcription antitermination where NusB:NusE is part of the antitermination complex[14,16,19]. The amount of free NusE that is not bound to the ribosome is estimated to be very low, but it is essential for transcription antitermination[36]. Thus tethering of NusE or the NusB:NusE complex to RNAP might be an early event in transcription antitermination to increase the local NusE concentration. NusE would remain bound to the TEC until transferred to NusG-CTD during assembly of the antitermination complex. As ribosomal operons comprise a very high density of transcribing RNAPs with high elongation rates[37], tethering NusE directly to RNAP would ensure fast and efficient transcription antitermination in these operons.

Solution structure of NusA-NTD from *E. coli*. The six domains comprising transcription factor NusA associates with RNAP *via* NusA-NTD, which is necessary and sufficient for the enhancement of pausing during transcription[27]. To determine the solution structure of NusA-NTD by NMR spectroscopy we initially tried a construct containing amino acids Met1-Ile137 carrying an N-terminal His$_9$-tag, NusA(1–137). The high degree of heterogeneity in the peak intensities as well as the spectral overlap in the [^{1}H,^{15}N]-HSQC spectrum of the [^{15}N]-labeled protein, however, prevented further analysis (Supplementary Fig. 3). A shorter construct, NusA-NTD$^{\Delta}$, consisting of amino acids Met1-Met125 and a cleavable C-terminal His$_6$-tag, led to homogeneous signal intensities with non-overlapping signals in the [^{1}H,^{15}N]-HSQC spectra (Supplementary Fig. 3) and allowed nearly complete backbone and side chain resonance assignment. No resonances were found for residues Asp103, Arg104, Thr106, Thr107, and Gln108. These are located in a flexible loop so that severe line broadening may occur caused by either fast solvent exchange or conformational exchange on the intermediate chemical shift time scale. Structure determination was performed on the basis of 1565 distance and 193 dihedral restraints derived from multiple NMR experiments (Table 1).

NusA-NTD$^{\Delta}$ comprises four α-helices (α1: Asn2–Ala17, α2: Pro19–Glu40, α3: Leu77–Glu85, α4: Thr106–Ala124) and four β-strands (β1: Val45–Asp50, β2: Asp55–Val65, β3: Glu74–Thr76, β4: Gly90–Gln96) and its structure resembles that of NusA-NTDs from other bacteria[22,28,38,39]. It is L-shaped, with a globular head and a mainly α-helical body (Fig. 5a and b). In the latter α1, α2, α4, β1, and β2 surround an elongated hydrophobic core, and the long β2 strand protrudes into the globular head. The C-terminal

Distance restraints	total	1507
	intraresidual	329
	sequential	386
	medium range	321
	long range	471
Hydrogen bond restraints		58
Dihedral restraints		193
Restraint violations	rms distance violation (Å)	0.006 (±0.0011)
	max. distance violation (Å)	0.11
	rms dihedral violation (°)	0.05 (±0.02)
	max. dihedral violation (°)	0.8
	rmsd bond length (Å)	0.00070 (±0.00009)
	rmsd bond angle (°)	0.13 (±0.012)
Atomic coordinate precision	backbone atoms (Å)	0.80[a]
	all heavy atoms (Å)	1.13[a]
Ramachandran plot statistics[b]	most favored regions (%)	90.5
	additional allowed regions (%)	8.8
	generously allowed regions (%)	0.2
	disallowed regions (%)	0.5

Table 1. Experimental constraints for structure calculation of NusA-NTD$^{\Delta}$. [a]residues Met1-Arg123. [b]determined by PROCHECK-NMR.

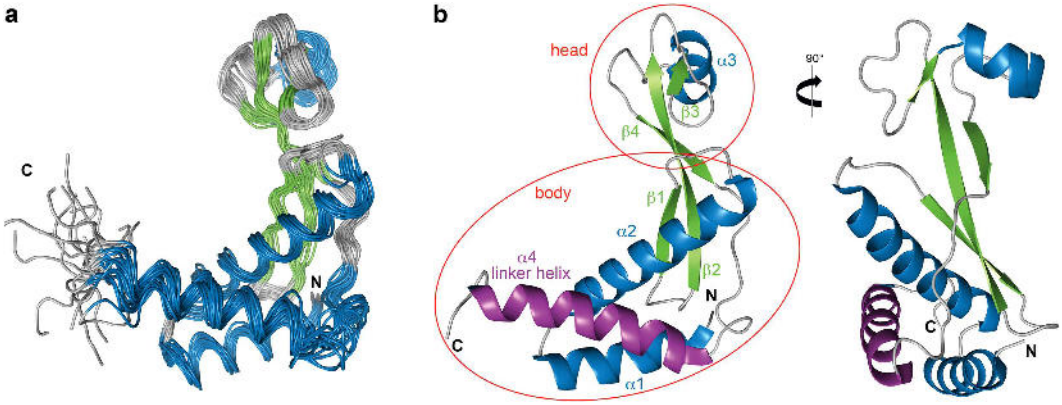


Figure 5. Solution structure of NusA-NTD$^{\Delta}$. (**a**) Structural ensemble of the 20 accepted lowest energy structures in ribbon representation colored according to secondary structure (α-helices, blue; β-strands, green; loops, grey). (**b**) Cartoon representation of the calculated structure with the lowest energy. Secondary structure elements are colored as in (**a**) and labeled. Helix α4 is highlighted in purple, the head and body parts are indicated.

helix α4 connects NusA-NTD and the NusA-SKK domain (linker helix). The globular head comprises α3, β3, β4, and the N-terminal part of β2. While the head is mainly acidic, the body exhibits large basic patches (Supplementary Fig. 4).

To date structures of NusA proteins from different bacteria are available, and although all NusA-NTDs are similar in their overall architecture, they differ in the position of the linker helix (Supplementary Fig. 5a–f). For NusA-NTD from *B. subtilis* (*Bs*NusA-NTD), NMR data suggest that this helix occurs in two alternative conformations in solution[28]. However, we have no indication for the presence of multiple conformations of helix α4 in NusA-NTD$^{\Delta}$. Moreover, unambiguous [^{15}N]-nuclear Overhauser enhancement spectroscopy (NOESY) cross peaks between hydrophobic amino acids could be observed in NMR experiments, demonstrating a direct interaction between helix α4 and helices α1 and α2 in NusA-NTD$^{\Delta}$ (Supplementary Fig. 5g). As crystal structures of full length NusA from *Thermotoga maritima* (*Tm*NusA, protein data bank (PDB) IDs: 1HH2, 2L2F), *Mycobacterium tuberculosis* (*Mt*NusA, PDB ID: 1K0R) and *Planctomyces limnophilus* (*Pl*NusA, PDB ID: 4MTN) show that the NusA-SKK domain is connected

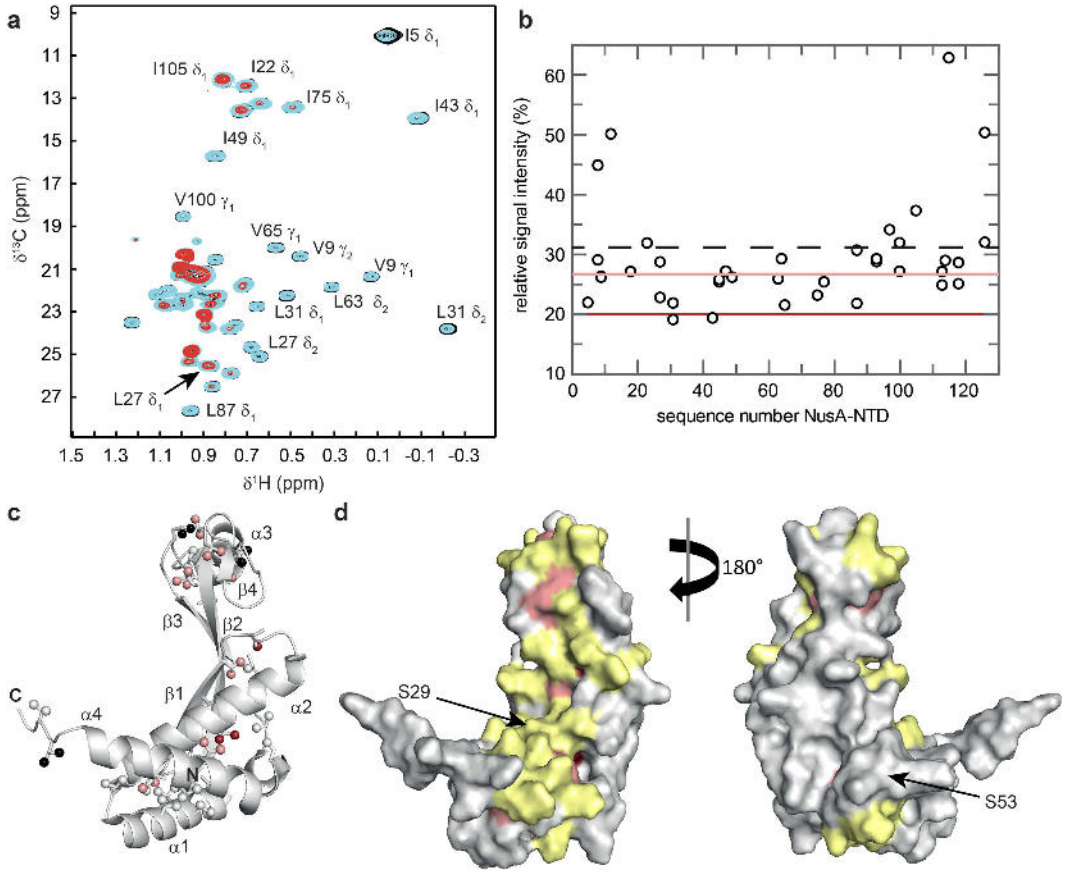


Figure 6. RNAP binding site of NusA-NTD$^{\Delta}$. (**a**) Titration of [I,L,V]-NusA-NTD$^{\Delta}$ with RNAP. Methyl-TROSY spectra of [I,L,V]-NusA-NTD$^{\Delta}$ in the absence, black, and in the presence of RNAP (1:1 molar ratio, cyan; 1:2 molar ratio, red), with assignment of representative signals. (**b**) Relative [I,L,V]-NusA-NTD$^{\Delta}$ signal intensity after addition of RNAP in equimolar concentration *vs*. amino acid sequence positions of NusA-NTD$^{\Delta}$. Dashed black line, average relative signal intensity; dark red and light red lines, thresholds for strongly affected (65% of the average relative intensity) and slightly affected (85% of the average relative intensity) residues, respectively. (**c**) Mapping of affected methyl groups onto the NusA-NTD$^{\Delta}$ structure. NusA-NTD$^{\Delta}$ (grey) in cartoon representation. Ile, Leu, and Val residues are in stick representation with the carbon atoms of their methyl groups as spheres. Strongly affected methyl groups, dark red; slightly affected methyl groups, light red; unaffected methyl groups, grey; unassigned methyl groups, black. (**d**) Mapping of affected residues onto the NusA-NTD$^{\Delta}$ structure (surface representation). For graphical illustration of the interaction site the complete amino acid was colored as affected in lieu of the methyl group. Colors are as in (**c**). Two amino acids on either side of an affected Ile/Leu/Val residue are highlighted in yellow unless they were unaffected Ile/Leu/Val residues. The positions of Ser29 and Ser53 are marked by black arrows.

to the linker helix by only a short loop, this helix might be responsible for the correct positioning of NusA-SKK for RNA binding.

Comparing NusA-NTD structures it is striking that *Mt*NusA-NTD and *Pl*NusA-NTD lack the globular head (Supplementary Fig. 5a–e), which is proposed to interact with the β′ subunit of RNAP[40]. This might indicate a different mode of action/binding of *Mt*NusA and *Pl*NusA compared to other NusAs.

RNAP interface of NusA-NTD. NusA-NTD is supposed to bind to RNAP by interacting with the β flap tip helix of the β flap region, which forms the outer wall of the RNA exit channel. To date, available complex models are based on a low-resolution electron microscopy structure, cleavage experiments, targeted amino acid exchanges and NMR experiments using a short β flap construct[26–28]. Here we used complete RNAP to determine the RNAP binding site of NusA-NTD$^{\Delta}$ by applying the same approach as for NusG-NTD and NusE$^{\Delta}$. Methyl group labeled NusA-NTD$^{\Delta}$ ([I,L,V]-NusA-NTD$^{\Delta}$) was titrated with protonated RNAP leading to a non-uniform decrease of [I,L,V]-NusA-NTD$^{\Delta}$ methyl group signals (Fig. 6a). Again, the normalized signal intensity decrease in the 1:1 complex was analyzed to identify highly and slightly affected methyl groups (Fig. 6b). These are located mainly on the concave side of the body and in the acidic head (Fig. 6c). Inspection of the surface representation suggests that the β-sheet on the concave side of NusA-NTD$^{\Delta}$ is the center of the interaction surface, although it contains only a

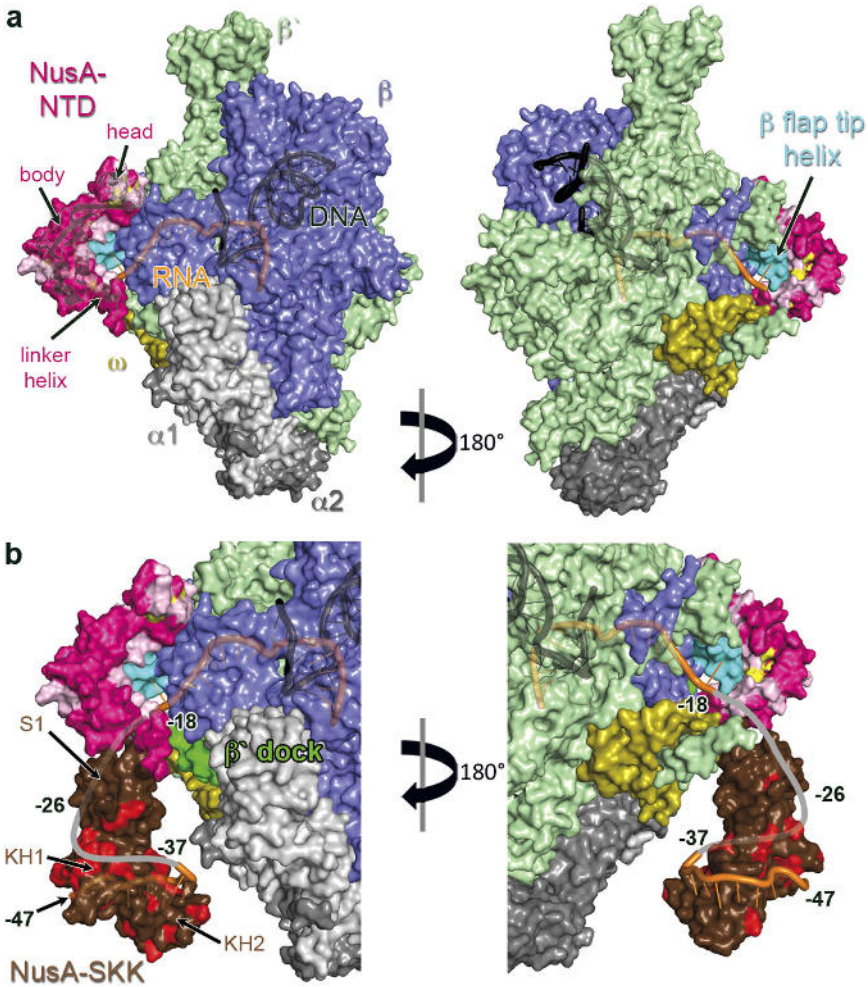


Figure 7. Model for the binding of NusA-NTD$^{\Delta}$ to elongating RNAP. (**a**) NusA-NTD$^{\Delta}$ (cartoon and surface representation, pink) is docked to elongating *Tt*RNAP (PDB ID: 2O5I, surface representation). Residues in NusA-NTD$^{\Delta}$ that are affected by RNAP binding are highlighted in yellow and two amino acids on either side of an affected Ile/Leu/Val residue are colored in light pink unless they were unaffected Ile/Leu/Val residues. α_1, light grey; α_2, dark grey; β, blue; β′, pale green; ω, olive; β flap tip helix, teal; RNA, orange; DNA, black. (**b**) Binding of exiting RNA by NusA. The orientation of NusA-NTD$^{\Delta}$ is the same as in (**a**), the position of *Tm*NusA-SKK was modeled by superposing *Tm*NusA-NTD (PDB ID: 1L2F) on NusA-NTD$^{\Delta}$. RNA was taken from the *Mt*NusA-SKK:RNA complex (PDB ID: 2ASB). Representation of NusA-NTD$^{\Delta}$, *Tt*RNAP and nucleic acids as in (**a**). The β′ dock domain is highlighted in green. *Tm*NusA-SKK (brown) is in surface representation with residues affected by RNA binding highlighted in red according to Schweimer *et al.*[4]. The grey line shows a possible path of exiting RNA, the estimated base numbers are indicated.

limited number of Ile, Leu, or Val residues resulting in a low structural resolution (Fig. 6d). Our binding site is in accordance with cleavage experiments using NusA variants NusA(S29C) and NusA(S53C), that indicated that S29 is located in the NusA:RNAP interface, while S53 is at the opposite side of NusA-NTD (Fig. 6d)[27]. Moreover, our results generally agree with mutational analyses showing that the concave side of the β-sheet is involved in NusA-NTD:β flap interaction[28].

Model of the NusA:RNAP complex. NusA has various effects on transcription elongation and termination with the NusA-NTD:RNAP interaction being probably one key step within the regulatory mechanism[27]. NusA-NTD contacts the RNA exit channel by binding to the β flap tip helix of the β flap region, but the resolution of the electron microscopy structure of a NusA-NTD:RNAP complex was too low to unambiguously determine the orientation of NusA-NTD bound to RNAP[26]. Cleavage and crosslinking experiments on the one hand and mutational analyses as well as NMR studies on *Bs*NusA-NTD and a short β flap construct on the other hand lead to two binding models[27,28].

We used our NMR data to dock NusA-NTD$^{\Delta}$ to the β flap tip helix of elongating *Thermus thermophilus* RNAP (*Tt*RNAP, PDB ID: 2O5I) using HADDOCK[41] (Fig. 7a). In the model most reliable according to HADDOCK, the body of NusA-NTD$^{\Delta}$ binds the β flap tip helix *via* its concave side, which

is in accordance with other models[27,28]. The body is oriented towards the RNA exit channel so that the globular head interacts with the β′ subunit, the latter being in agreement with previous findings that the β′ subunit might also be involved in NusA-NTD binding[20,40]. This orientation allows a tight interaction with the *Tt*RNAP and is similar to the orientations suggested in earlier models[27,28], although the absolute position of NusA-NTD$^{\Delta}$ strongly depends on the residues chosen as restraints and the position of the β flap tip helix.

Next, we integrated the NusA-SKK domain into the model (Fig. 7b). As the structure of *E. coli* NusA-SKK is not available and as the position of the linker helix is similar in *Pl*NusA and NusA-NTD$^{\Delta}$, we first used the crystal structure of *Pl*NusA as template. This, however, led to heavy steric clashes of the *Pl*NusA-SKK domain and *Tt*RNAP which could be prevented by rotating the *Pl*NusA-SKK domain away from the *Tt*RNAP, using the 3-4 residues following the linker helix as anchor. Alternatively, the linker helix itself might rotate slightly. Thus, we modeled the position of *Tm*NusA-SKK by superposing *Tm*NusA-NTD (PDB ID: 1L2F) on NusA-NTD$^{\Delta}$, and we added a short piece of RNA from the *Mt*NusA-SKK:RNA complex structure (PDB ID: 2ASB, Fig. 7b). Either way, the NusA-SKK domain can be positioned correctly for RNA binding. As NusA-NTD is necessary and sufficient for enhancing transcriptional pausing and recognizes duplex RNA[27], exiting RNA might first contact a basic patch on the helical bundle of the NusA-NTD body (Supplementary Fig. 4), which is in direct vicinity of the RNAP exit channel. The RNA then wraps around the NusA-SKK domain, which, in turn, recognizes specific RNA signals (Fig. 7b)[4,42,43]. Crosslinking experiments showed that the RNA region −16 to −23 lies near the NusA-NTD in full-length NusA and that the −34 to −40 region of exiting RNA contacts the NusA-KH2 domain[27], which is consistent with our model. Moreover, the NusA-S1 domain is placed in the vicinity of the β′ dock domain, being in accordance with a genetically shown NusA-S1:β′ dock interaction[44] and cleavage experiments using Fe(III)-(*S*)-2-[4-(2-bromoacetamido)benzyl]ethylenediaminetetraacetic acid (FeBABE)[27]. The position of the C-terminus of NusA-SKK roughly orientates the two NusA-AR domains towards the α-subunits of RNAP and thus localizes NusA-AR2 close to the α-CTD, sterically simplifying a NusA-AR2:α-CTD interaction[4].

Finally, it has been speculated that reorientation of helix α4 stabilizes RNA hairpins[28]. However, not only does NusA exhibit large conformational plasticity, but, in addition, the β flap tip helix is also a highly mobile element[28]. During the transcription cycle the flexibility of the β flap tip helix is important for the regulation of the size of the RNA exit channel, of which the β flap forms the outer wall. Thus, we suggest that the orientation of NusA-NTD bound to RNAP as well as the position of helix α4 may vary, depending on the position of the β flap tip helix. Moreover, this structural flexibility is complemented by the other NusA domains, which are all elastically connected.

Outlook. In this conceptually simple single-experiment approach to identify the RNAP interaction surface of transcription factors with NMR spectroscopy (i) complete RNAP is used, (ii) probes in the transcription factor are directly monitored and, most importantly, (iii) none of the interaction partners needs to be modified. In the future, the method will be refined and used to study these interactions in more detail. Moreover, this approach is very general and can thus be transferred to other systems, with a small binding partner interacting with a supramolecular complex.

Materials and Methods

Cloning. The gene coding for *Ec*NusA-NTD(1–137) was cloned into pET19b *via Blp*I and *Bam*HI. The resulting *E. coli* expression vector pET19b_NusA-NTD_1-137 codes for a His_9 tag fused to the N-terminus of NusA-NTD, cleavable by PreScission protease.

Gene expression and protein purification. NusG-NTD was produced and purified as described[45], as was NusA-NTD$^{\Delta}$[20], the NusB:NusE$^{\Delta}$ complex[34,46] and RNAP[20].

Expression of *nusA-NTD(1–137)* was carried out in *E. coli* BL21 (λ DE3) (Novagen, Madison, WI, USA) harboring pET19b_NusA-NTD_1-137. Lysogeny broth (LB) medium supplemented with 100 μg/ml ampicillin was inoculated with a preculture to an optical density at 600 nm (OD_{600}) of 0.2 and cells were grown at 37 °C until they reached an OD_{600} of 0.7. The temperature was lowered to 20 °C and 30 min later overexpression was induced with 2 mM IPTG. After overnight growth, cells were harvested by centrifugation (9,000 x *g*, 15 min, 4 °C) and dissolved in 20 mM tris(hydroxymethyl)aminomethane (Tris)/HCl (pH 7.9), 100 mM NaCl, 10% (v/v) glycerol, 5 mM β-mercaptoethanol, 10 mM imidazole (buffer A). Cell disruption was carried out with a microfluidizer (Microfluidics, Newton, MA, USA). Having centrifuged the lysate (12,000 × *g*, 30 min, 4 °C), the supernatant was applied to a Ni-NTA column (Qiagen, Hilden, Germany), and subsequently the column was washed with buffer A. A step gradient with increasing imidazole concentrations (10–500 mM in buffer A) was used for elution. Fractions containing His_9-NusA-NTD(1–137) were combined and cleaved during overnight dialysis against 50 mM Tris/HCl (pH 8.0), 150 mM NaCl (molecular weight cut-off (MWCO) 3,500 Da) by PreScission protease (GE Healthcare, Munich, Germany). The protein solution was then dialyzed against 50 mM Tris (pH 7.4), 1 mM dithiothreitol (DTT, buffer B) and reapplied to the Ni-NTA column connected to a QXL FF column (GE Healthcare, Munich, Germany). After washing with buffer B, the Ni-NTA column was removed and the QXL FF column was eluted using a step gradient with increasing NaCl concentrations (0–1 M NaCl in buffer B). Fractions containing pure NusA-NTD(1-137) were dialyzed against the

required buffer, concentrated by ultrafiltration (MWCO 3,000 Da) and stored at −80 °C after freezing with liquid nitrogen.

Proteins were uniformly labeled with ^{15}N or $^{15}N,^{13}C$ by growing *E. coli* in M9 minimal medium[41,42] with addition of $(^{15}NH_4)_2SO_4$ (Campro Scientific, Berlin, Germany) or $(^{15}NH_4)_2SO_4$ and ^{13}C-D-glucose (Spectra Stable Isotopes, Columbia, MD, USA) as only nitrogen and carbon source. Expression and purification was the same as for proteins produced in LB medium. Methyl group labeling of Ile, Leu and Val residues with [$^{1}H,^{13}C$] in deuterated proteins was performed as described previously[20].

NMR spectroscopy. NMR spectroscopic experiments were conducted on Bruker *Avance* 600 MHz, 700 MHz and 800 MHz spectrometers, the latter two equipped with cryogenically cooled probes. For resonance assignment of NusA-NTD$^{\Delta}$, standard double and triple resonance through-bond experiments were recorded[47,48]. The protein was in 10 mM potassium phosphate buffer (pH 6.4) containing 50 mM NaCl at 298 K. NMR data were processed using in-house routines (Apodization, Fourier transformation, phase correction and baseline correction) and visualized with NMRView[49]. Distance restraints for structure calculation were derived from [^{15}N]-edited and [^{13}C]-edited NOESY spectra with mixing times of 100–120 ms. NOESY cross peaks were classified according to their relative intensities and converted to distance restraints with the following upper limits: 3.0 Å, strong; 4.0 Å, medium; 5.0 Å, weak; 6.0 Å, very weak. Experimental NOESY spectra were validated semi-quantitatively against back-calculated spectra to confirm the assignment and to avoid bias of upper distance restraints by spin-diffusion. Hydrogen bonds were included for backbone amide protons in regular secondary structure if the amide proton did not show a water exchange cross peak in the [^{15}N]-edited NOESY spectrum. Backbone dihedral restraints were obtained from chemical shift data by using TALOS[50]. Existence of a hydrogen bond was assumed if the acceptor of a slowly exchanging amide proton, based on the absence of a water exchange peak in the [^{15}N]-edited NOESY spectrum, could be identified unambiguously from the results of initial structure calculations. For each hydrogen bond the distance between the amide proton and the acceptor was restrained to less than 2.3 Å and the distance between the amide nitrogen and the acceptor to less than 3.1 Å.

The structure calculation was performed with the program XPLOR-NIH 2.1.2[51] using a three-step simulated annealing protocol with floating assignment of prochiral groups including a conformational database potential[52]. For the final iteration 80 structures were calculated, the 20 structures of lowest energy were accepted and further analyzed with the programs XPLOR-NIH 2.1.2 and PROCHECK-NMR[53].

TROSY spectra[29] were recorded using [I,L,V]-labeled protein samples (20 μM) in 25 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES, pH 7.5), 50 mM NaCl, 5% (v/v) glycerol, 0.5 mM ethylenediaminetetraacetic acid (EDTA), 10 mM $MgCl_2$, 10 μM $ZnCl_2$, 1 mM DTT in 99.9% D_2O at 298 K. Unlabeled, protonated RNAP in the same buffer was added in two steps (ratios 1:1, 1:2). Non-stereo-specific assignments of methyl groups of NusG-NTD and NusE$^{\Delta}$ were taken from previous studies[10,46]. Signal intensities were normalized by protein concentration and number of scans. As pulse lengths changed less than 1% upon RNAP addition, the influence of these changes on the intensity were neglected. For each titration step the ratio of remaining signal intensities and signal intensities in the spectrum of the free transcription factor were calculated, yielding relative signal intensities. Next, the mean value of all relative intensities in each titration step was determined and experiment-specific thresholds of the mean value were defined. Residues with relative signal intensities below these thresholds were classified as either strongly or slightly affected. Additionally, Leu and Val residues were considered as affected, when at least one of the two signals showed a significant intensity decrease. Only unambiguously assigned signals were used in the analysis.

Proteins for the displacement experiments of [^{15}N]-NusE$^{\Delta}$:NusB from RNAP by NusG-CTD and of NusE$^{\Delta}$:NusB from [^{15}N]-NusG-CTD by RNAP were in 25 mM HEPES, pH 7.5, 100 mM NaCl at 298 K. Separate samples for [^{15}N]-NusE$^{\Delta}$:NusB (50 μM) and [^{15}N]-NusE$^{\Delta}$:NusB:RNAP (25 μM each) were prepared. For the displacement experiments NusG-CTD was added (stock concentration: 1050 μM). Similarly, separate samples for [^{15}N]-NusG-CTD (50 μM) and [^{15}N]-NusG-CTD: NusE$^{\Delta}$:NusB (25 μM each) were prepared. For the displacement experiments RNAP was added from a 117 μM stock. The titrations were followed by recording 1D or 2D [$^{1}H,^{15}N$]-HSQC spectra after each titration step. 1D spectra were normalized by protein concentration and number of scans. As pulse lengths changed less than 1% upon RNAP addition, the influence of these changes on the intensity were neglected.

Docking and Molecular Modeling. The NusG-NTD:RNAP complex was generated based on the crystal structure of Spt4/5 bound to the clamp domain from *P. furiosus* (PDB ID: 3QQC). *E. coli* NusG-NTD (PDB ID: 2K06, model 1) was superposed on Spt5 (chain D, root mean square deviation (r.m.s.d.) 1.2 Å). *Ec*RNAP (PDB ID: 4KMU) was positioned by superposing the β′ subunit (chain D) on the clamp domain (chain A, r.m.s.d. 2.4 Å).

Docking of NusA-NTD$^{\Delta}$ (model 1) to elongating *Tt*RNAP (PDB ID: 2O5I) was carried out using the HADDOCK webserver[41]. Residues in NusA-NTD$^{\Delta}$ that were experimentally determined to be affected by RNAP binding (Leu27, Leu31, Ile43, Val45) were defined as active residues. Solvent exposed residues in the β flap tip helix were chosen as active residues (chain C, residues Arg772, Leu773, Ser776, Ile777). Passive residues were automatically determined by HADDOCK. The coordinates of the β flap tip helix in the docked complex relative to the deposited coordinates of NusA-NTD$^{\Delta}$ are shown in Supplementary

Table 1. After docking NusA-NTD$^{\Delta}$ to *Tt*RNAP, the position of the NusA-SKK domain was modeled with two alternative procedures. First, *Pl*NusA (PDB ID: 4MTN) was superposed on NusA-NTD$^{\Delta}$ (residues G3-D73 of *Pl*NusA; residues Met1-Thr101 of NusA-NTD$^{\Delta}$). To avoid clashes with *Tt*RNAP the *Pl*NusA-SKK was rotated manually around residues in the linker between *Pl*NusA-NTD and *Pl*NusA-SKK (residues Arg107-Gln109) using PyMOL[54]. In the second approach *Tm*NusA (PDB ID: 1L2F) was superposed on NusA-NTD$^{\Delta}$ using residues 1–101. Finally, the *Mt*NusA-SKK:RNA complex (PDB ID: 2ASB, residues Ser108-Gly333 of *Mt*NusA-SKK) was superposed on *Tm*NusA-SKK (residues Glu132-Leu344) to position the RNA. RNA base numbers were estimated.

Programs. All structures were visualized with PyMOL[54]. The Adaptive Poisson-Boltzmann Solver (APBS)-Plugin and the PDB2PQR server were used for the determination of the charge surface potential[55,56]. Superpositions of different NusA-NTDs were done with LSQMAN[57], omitting the linker helix (residues Met1-Thr101 of NusA-NTD$^{\Delta}$, residues Met1-Asn101 of *Tm*NusA (PDB ID: 1L2F, 1HH2), residues Met1-Asp101 of *Bs*NusA (PDB ID: 2MT4), residues Met1-Phe79 of *Mt*NusA (PDB ID: 2K0R), residues Gly3-Asp73 of *Pl*NusA (PDB ID: 4MTN)). All other superpositions were carried out by PyMOL[54].

References

1. Werner, F. & Grohmann, D. Evolution of multisubunit RNA polymerases in the three domains of life. *Nat. Rev. Microbiol.* **9,** 85–98 (2011).
2. Ito, K., Iwakura, Y. & Ishihama, A. Biosynthesis of RNA polymerase in *Escherichia coli*. III. Identification of intermediates in the assembly of RNA polymerase. *J. Mol. Biol.* **96,** 257–271 (1975).
3. Jeon, Y. H., Yamazaki, T., Otomo, T., Ishihama, A. & Kyogoku, Y. Flexible linker in the RNA polymerase alpha subunit facilitates the independent motion of the C-terminal activator contact domain. *J. Mol. Biol.* **267,** 953–962 (1997).
4. Schweimer, K. *et al.* NusA interaction with the α-subunit of *E. coli* RNA Polymerase is *via* the UP-element site and releases autoinhibition. *Structure* **19,** 945–954 (2011).
5. Mustaev, A. *et al.* Modular organization of the catalytic center of RNA polymerase. *Proc. Natl. Acad. Sci. USA.* **94,** 6641–6645 (1997).
6. Zaychikov, E. *et al.* Mapping of catalytic residues in the RNA polymerase active center. *Science* **273,** 107–109 (1996).
7. Ghosh, P., Ishihama, A. & Chatterji, D. *Escherichia coli* RNA polymerase subunit omega and its N-terminal domain bind full-length beta′ to facilitate incorporation into the alpha2beta subassembly. *Eur. J. Biochem.* **268,** 4621–4627 (2001).
8. Haugen, S. P., Ross, W. & Gourse, R. L. Advances in bacterial promoter recognition and its control by factors that do not bind DNA. *Nat. Rev. Microbiol.* **6,** 507–519 (2008).
9. Werner, F. A nexus for gene expression-molecular mechanisms of Spt5 and NusG in the three domains of life. *J. Mol. Biol.* **417,** 13–27 (2012).
10. Mooney, R. A., Schweimer, K., Rösch, P., Gottesman, M. E. & Landick, R. 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 (2009).
11. Artsimovitch, I. & Landick, R. Pausing by bacterial RNA polymerase is mediated by mechanistically distinct classes of signals. *Proc. Natl. Acad. Sci. USA.* **97,** 7090–7095 (2000).
12. Martinez-Rucobo, F. W., Sainsbury, S., Cheung, A. C. & Cramer, P. Architecture of the RNA polymerase-Spt4/5 complex and basis of universal transcription processivity. *EMBO J.* **30,** 1302–1310 (2011).
13. Sevostyanova, A., Belogurov, G. A., Mooney, R. A., Landick, R. & Artsimovitch, I. The β subunit gate loop is required for RNA polymerase modification by RfaH and NusG. *Mol. Cell* **43,** 253–262 (2011).
14. Burmann, B. M. *et al.* A NusE:NusG Complex Links Transcription and Translation. *Science* **328,** 501–504 (2010).
15. Li, J., Mason, S. W. & Greenblatt, J. Elongation factor NusG interacts with termination factor rho to regulate termination and antitermination of transcription. *Genes Dev.* **7,** 161–172 (1993).
16. Mason, S. W. & Greenblatt, J. 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 (1991).
17. Aksoy, S., Squires, C. L. & Squires, C. Evidence for antitermination in *Escherichia coli* RRNA transcription. *J. Bacteriol.* **159,** 260–4 (1984).
18. Gottesman, M. E. & Weisberg, R. A. Little lambda, who made thee? *Microbiol. Mol. Biol. Rev.* **68,** 796–813 (2004).
19. Nodwell, J. R. & Greenblatt, J. Recognition of *boxA* antiterminator RNA by the *E. coli* antitermination factors NusB and ribosomal protein S10. *Cell* **72,** 261–268 (1993).
20. Drögemüller, J. *et al.* Exploring RNA polymerase regulation by NMR spectroscopy. *Sci. Rep.* **5,** 10825–10835 (2015).
21. Eisenmann, A., Schwarz, S., Prasch, S., Schweimer, K. & Rösch, P. The *E. coli* NusA carboxy-terminal domains are structurally similar and show specific RNAP- and lambdaN interaction. *Protein Sci.* **14,** 2018–29 (2005).
22. Worbs, M., Bourenkov, G. P., Bartunik, H. D., Huber, R. & Wahl, M. C. An extended RNA binding surface through arrayed S1 and KH domains in transcription factor NusA. *Mol. Cell* **7,** 1177–1189 (2001).
23. Borukhov, S., Lee, J. & Laptenko, O. Bacterial transcription elongation factors: new insights into molecular mechanism of action. *Mol. Microbiol.* **55,** 1315–1324 (2005).
24. Roberts, J. W., Shankar, S. & Filter, J. J. RNA polymerase elongation factors. *Annu. Rev. Microbiol.* **62,** 211–233 (2008).
25. Mah, T. F., Li, J., Davidson, A. R. & Greenblatt, J. 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 (1999).
26. Yang, X. *et al.* The structure of bacterial RNA polymerase in complex with the essential transcription elongation factor NusA. *EMBO Rep.* **10,** 997–1002 (2009).
27. Ha, K. S., Toulokhonov, I., Vassylyev, D. G. & Landick, R. The NusA N-terminal domain is necessary and sufficient for enhancement of transcriptional pausing *via* interaction with the RNA exit channel of RNA polymerase. *J. Mol. Biol.* **401,** 708–725 (2010).
28. Ma, C. *et al.* RNA polymerase-induced remodelling of NusA produces a pause enhancement complex. *Nucl. Acids Res.* **43,** 2829–2840 (2015).
29. Tugarinov, V. & Kay, L. E. An isotope labeling strategy for methyl TROSY spectroscopy. *J. Biomol. NMR* **28,** 165–172 (2004).
30. Epshtein, V., Dutta, D., Wade, J. & Nudler, E. An allosteric mechanism of Rho-dependent transcription termination. *Nature* **463,** 245–249 (2010).
31. Gnatt, A. L., Cramer, P., Fu, J., Bushnell, D. A. & Kornberg, R. D. Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 Å resolution. *Science* **292,** 1876–1882 (2001).

32. Weixlbaumer, A., Leon, K., Landick, R. & Darst, S. A. Structural basis of transcriptional pausing in bacteria. *Cell* **152,** 431–441 (2013).
33. Davis, C. A., Bingman, C. A., Landick, R., Record, M. T. Jr & Saecker, R. M. Real-time footprinting of DNA in the first kinetically significant intermediate in open complex formation by *Escherichia coli* RNA polymerase. *Proc. Natl. Acad. Sci. USA*. **104,** 7833–7838 (2007).
34. Luo, X. *et al.* Structural and functional analysis of the *E. coli* NusB-S10 transcription antitermination complex. *Mol. Cell* **32,** 791–802 (2008).
35. Schuwirth, B. S. *et al.* Structures of the Bacterial Ribosome at 3.5 A Resolution. *Science* **310,** 827–834 (2005).
36. Greive, S. J., Lins, A. F. & von Hippel, P. H. 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 (2005).
37. Vogel, U. & Jensen, K. F. The RNA chain elongation rate in *Escherichia coli* depends on the growth rate. *J. Bacteriol*. **176,** 2807–2813 (1994).
38. Shin, D. H. *et al.* Crystal structure of NusA from *Thermotoga maritima* and functional implication of the N-terminal domain. *Biochemistry* **42,** 13429–13437 (2003).
39. Gopal, B. *et al.* Crystal structure of the transcription elongation/anti-termination factor NusA from *Mycobacterium tuberculosis* at 1.7 A resolution. *J. Mol. Biol.* **314,** 1087–1095 (2001).
40. Traviglia, S. L., Datwyler, S. A., Yan, D., Ishihama, A. & Meares, C. F. Targeted protein footprinting: where different transcription factors bind to RNA polymerase. *Biochemistry* **38,** 15774–15778 (1999).
41. de Vries, S. J., van Dijk, M. & Bonvin, A. M. The HADDOCK web server for data-driven biomolecular docking. *Nat. Protoc.* **5,** 883–897 (2010).
42. Beuth, B., Pennell, S., Arnvig, K. B., Martin, S. R. & Taylor, I. A. Structure of a *Mycobacterium tuberculosis* NusA-RNA complex. *EMBO J.* **24,** 3576–3587 (2005).
43. Prasch, S. *et al.* RNA-binding specificity of *E. coli* NusA. *Nucl. Acids Res*. **37,** 4736–4742 (2009).
44. Ito, K. & Nakamura, Y. Localization of *nusA*-suppressing amino acid substitutions in the conserved regions of the beta' subunit of *Escherichia coli* RNA polymerase. *Mol. Gen. Genet.* **251,** 699–706 (1996).
45. Burmann, B. M., Scheckenhofer, U., Schweimer, K. & Rösch, P. Domain interactions of the transcription-translation coupling factor *Escherichia coli* NusG are intermolecular and transient. *Biochem. J.* **435,** 783–789 (2011).
46. Burmann, B. M., Luo, X., Wahl, M. C., Rösch, P. & Gottesman, M. E. Fine tuning of the *E. coli* NusB:NusE complex affinity to *BoxA* RNA is required for processive antitermination. *Nucl. Acids Res*. **38,** 314–326 (2010).
47. Sattler, M., Schleucher, J. & Griesinger, C. Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. *Prog. NMR Spectrosc*. **34,** 93–158 (1999).
48. Cavanagh, J., Fairbrother, W. J., Palmer III, A. G., Rance, M. & Skelton, N. J. in Protein NMR spectroscopy: principles and practice (Academic Press, Boston, 2007).
49. Johnson, B. A. Using NMRView to visualize and analyze the NMR spectra of macromolecules. *Methods Mol. Biol.* **278,** 313–352 (2004).
50. Cornilescu, G., Delaglio, F. & Bax, A. Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *J. Biomol. NMR* **13,** 289–302 (1999).
51. Schwieters, C. D., Kuszewski, J. J., Tjandra, N. & Clore, G. M. The Xplor-NIH NMR molecular structure determination package. *J. Magn. Reson.* **160,** 66–74 (2003).
52. Schweimer, K. *et al.* Structural investigation of the binding of a herpesviral protein to the SH3 domain of tyrosine kinase Lck. *Biochemistry* **41,** 5120–5130 (2002).
53. Laskowski, R. A., Rullmann, J. A. C., MacArthur, M. W., Kaptein, R. & Thornton, J. M. AQUA and PROCHECK-NMR: Programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR* **8,** 477–486 (1996).
54. Schrödinger L. The PyMOL molecular graphics system, version 1.3. *Schrödinger, LLC, Mannheim, Germany* (2010).
55. Baker, N. A., Sept, D., Joseph, S., Holst, M. J. & McCammon, J. A. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. USA*. **98,** 10037–10041 (2001).
56. Dolinsky, T. J., Nielsen, J. E., McCammon, J. A. & Baker, N. A. PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucl. Acids Res*. **32,** W665–7 (2004).
57. Kleywegt, G. J. Use of non-crystallographic symmetry in protein structure refinement. *Acta Crystallogr. D Biol. Crystallogr*. **52,** 842–857 (1996).

Acknowledgements

We thank Ramona Heissmann for excellent technical assistance and Björn M. Burmann as well as Stefan Prasch for initial work on NusA-NTD and NusA-NTD$^{\Delta}$. The work was supported by grants Ro 617/17-1 and Ro 617/21-1 (to P.R.) from the Deutsche Forschungsgemeinschaft, and the Ludwig-Schaefer-Scholarschip 2015 from Columbia University Medical Center (to P.R.). We also thank Björn M. Burmann for carefully reading the manuscript and helpful discussions.

Author Contributions

P.R. and S.H.K. conceived and designed the project. M.S., J.D. and S.H.K. designed the NMR experiments. J.D. and M.S. conducted the experiments. M.J. and K.S. solved the NusA-NTD structure. J.D., M.S., K.S., S.H.K. and P.R. wrote the manuscript.

Additional Information

Accession codes: Chemical shifts of NusA-NTD$^{\Delta}$ have been deposited in the Biological Magnetic Resonance Bank Databank, accession number 16868. The atomic coordinates of the NusA-NTD$^{\Delta}$ structure have been deposited in the Protein Data Bank, accession number 2KWP.

Supplementary information accompanies this paper at http://www.nature.com/srep

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Drögemüller, J. *et al.* Determination of RNA polymerase binding surfaces of transcription factors by NMR spectroscopy. *Sci. Rep.* **5**, 16428; doi: 10.1038/srep16428 (2015).

Supplementary Information

Determination of RNA polymerase binding surfaces of transcription factors by NMR spectroscopy

Johanna Drögemüller[1,§], Martin Strauß[1,§], Kristian Schweimer[1], Marcel Jurk[2], Paul Rösch[1], Stefan H Knauer[1,*]

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

[2] present address: Max Planck Institute for Molecular Genetics, Ihnestr. 63-73, 14195 Berlin, Germany

[§] These authors contributed equally to this work

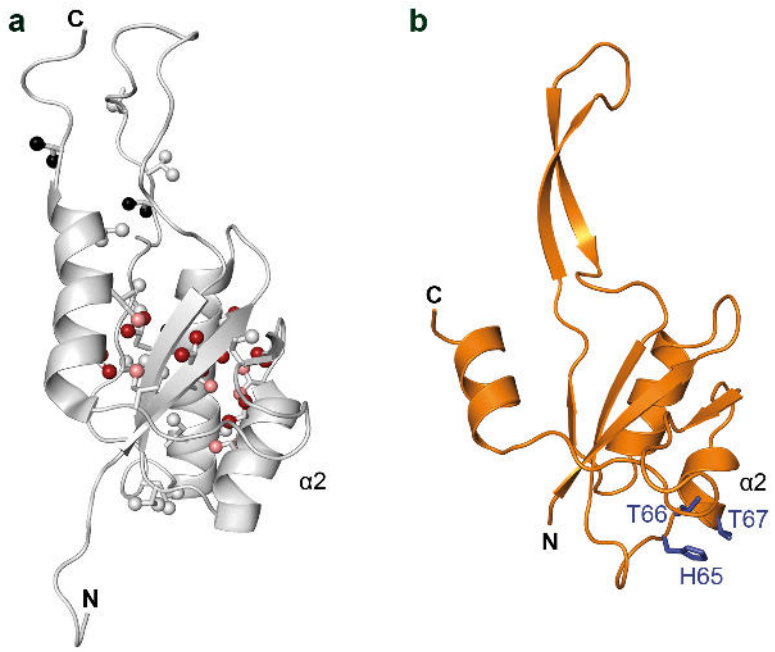
* corresponding author

Contents

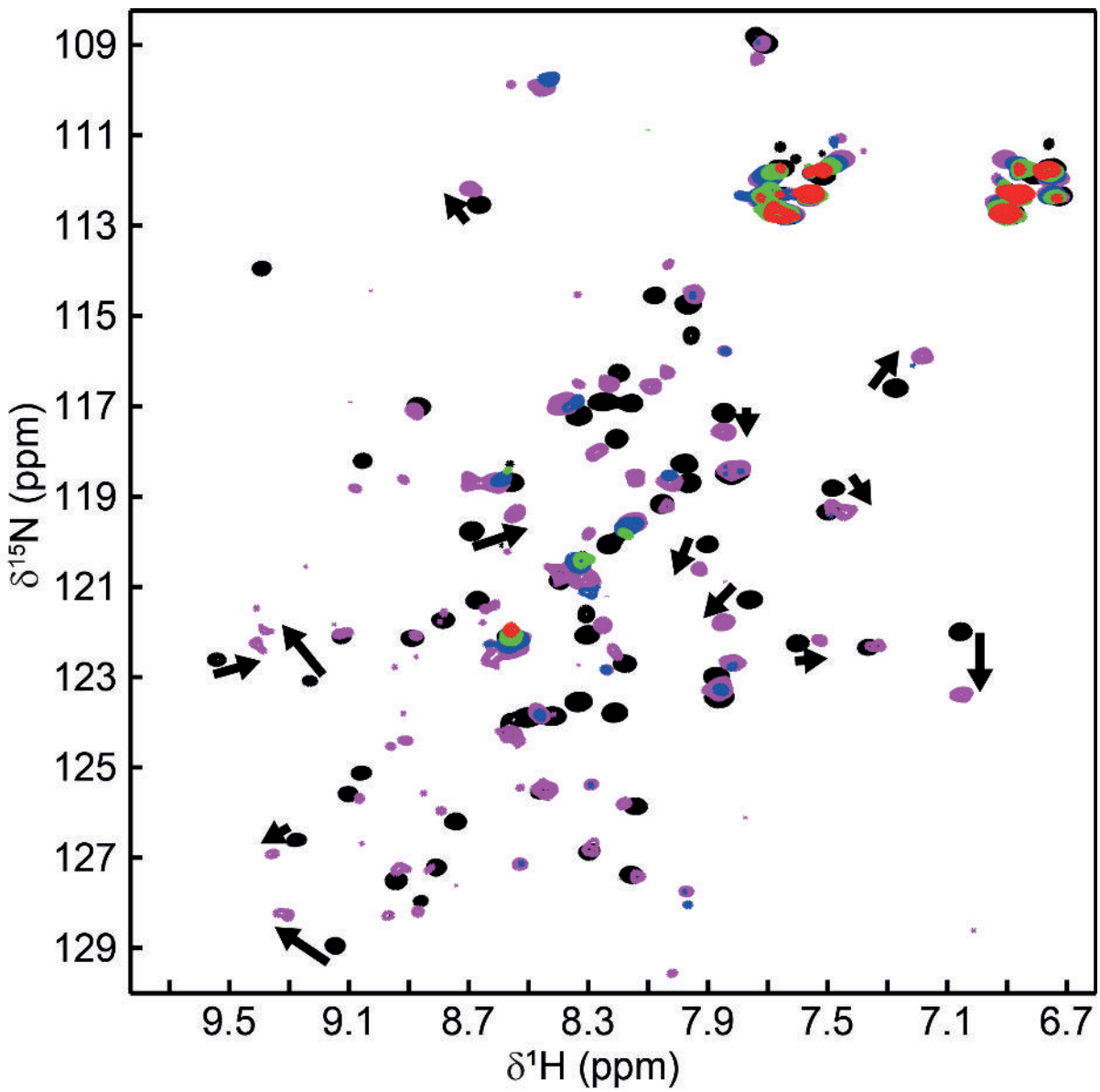
Supplementary Table 1: Coordinates of the β flap tip helix in the modeled NusA-NTD$^{\Delta}$:*Tt*RNAP complex. The table is an extract of the PDB file of elongating *Tt*RNAP (residues 767-781 of the β subunit; PDB ID: 2O5I) docked to NusA-NTD$^{\Delta}$ as described in the Material and Methods section, giving the position of *Tt*RNAP relative to the deposited coordinates of NusA-NTD$^{\Delta}$ (PDB ID: 2KWP).

ATOM	10777	N	PRO	C	767	3.453	-1.849	-42.571	1.00	90.52	N
ATOM	10778	CA	PRO	C	767	2.398	-2.073	-41.584	1.00	91.06	C
ATOM	10779	C	PRO	C	767	2.633	-1.165	-40.362	1.00	92.21	C
ATOM	10780	O	PRO	C	767	3.412	-0.213	-40.472	1.00	92.68	O
ATOM	10781	CB	PRO	C	767	2.524	-3.559	-41.274	1.00	90.38	C
ATOM	10782	CG	PRO	C	767	3.028	-4.129	-42.533	1.00	90.03	C
ATOM	10783	CD	PRO	C	767	4.098	-3.134	-42.897	1.00	90.44	C
ATOM	10784	N	THR	C	768	2.013	-1.418	-39.212	1.00	92.37	N
ATOM	10785	CA	THR	C	768	2.261	-0.517	-38.082	1.00	93.42	C
ATOM	10786	C	THR	C	768	2.739	-1.281	-36.849	1.00	94.34	C
ATOM	10787	O	THR	C	768	2.104	-2.254	-36.442	1.00	94.97	O
ATOM	10788	CB	THR	C	768	0.994	0.286	-37.704	1.00	92.74	C
ATOM	10789	OG1	THR	C	768	0.329	0.726	-38.894	1.00	92.51	O
ATOM	10790	CG2	THR	C	768	1.369	1.533	-36.866	1.00	91.14	C
ATOM	10791	N	PRO	C	769	3.867	-0.844	-36.239	1.00	95.40	N
ATOM	10792	CA	PRO	C	769	4.475	-1.458	-35.040	1.00	95.44	C
ATOM	10793	C	PRO	C	769	3.577	-1.368	-33.769	1.00	95.51	C
ATOM	10794	O	PRO	C	769	3.583	-2.294	-32.943	1.00	95.59	O
ATOM	10795	CB	PRO	C	769	5.803	-0.706	-34.892	1.00	96.45	C
ATOM	10796	CG	PRO	C	769	6.147	-0.342	-36.332	1.00	95.78	C
ATOM	10797	CD	PRO	C	769	4.806	0.125	-36.849	1.00	95.78	C
ATOM	10798	N	GLU	C	770	2.827	-0.273	-33.609	1.00	93.98	N
ATOM	10799	CA	GLU	C	770	1.962	-0.077	-32.434	1.00	93.34	C
ATOM	10800	C	GLU	C	770	0.563	-0.681	-32.666	1.00	93.16	C
ATOM	10801	O	GLU	C	770	-0.118	-1.081	-31.718	1.00	92.30	O
ATOM	10802	CB	GLU	C	770	1.849	1.427	-32.127	1.00	92.37	C
ATOM	10803	CG	GLU	C	770	3.152	2.249	-32.385	1.00	90.19	C
ATOM	10804	CD	GLU	C	770	3.957	2.594	-31.128	1.00	88.39	C
ATOM	10805	OE1	GLU	C	770	4.437	1.679	-30.437	1.00	87.21	O
ATOM	10806	OE2	GLU	C	770	4.113	3.798	-30.832	1.00	87.50	O
ATOM	10807	N	GLU	C	771	0.159	-0.730	-33.938	1.00	93.66	N
ATOM	10808	CA	GLU	C	771	-1.116	-1.314	-34.376	1.00	95.22	C
ATOM	10809	C	GLU	C	771	-0.991	-2.838	-34.494	1.00	96.29	C
ATOM	10810	O	GLU	C	771	-1.993	-3.557	-34.560	1.00	96.81	O
ATOM	10811	CB	GLU	C	771	-1.531	-0.719	-35.743	1.00	95.59	C
ATOM	10812	CG	GLU	C	771	-2.798	-1.302	-36.423	1.00	95.80	C
ATOM	10813	CD	GLU	C	771	-2.986	-0.867	-37.893	1.00	96.05	C
ATOM	10814	OE1	GLU	C	771	-4.129	-0.944	-38.403	1.00	95.71	O
ATOM	10815	OE2	GLU	C	771	-1.999	-0.457	-38.544	1.00	95.66	O
ATOM	10816	N	ARG	C	772	0.255	-3.318	-34.526	1.00	95.94	N
ATOM	10817	CA	ARG	C	772	0.562	-4.755	-34.598	1.00	96.21	C
ATOM	10818	C	ARG	C	772	0.310	-5.431	-33.247	1.00	97.11	C
ATOM	10819	O	ARG	C	772	-0.155	-6.577	-33.193	1.00	97.59	O
ATOM	10820	CB	ARG	C	772	2.031	-4.975	-34.992	1.00	94.95	C
ATOM	10821	CG	ARG	C	772	2.497	-6.441	-35.071	1.00	92.48	C
ATOM	10822	CD	ARG	C	772	1.926	-7.149	-36.290	1.00	91.56	C
ATOM	10823	NE	ARG	C	772	1.943	-6.309	-37.490	1.00	90.45	N
ATOM	10824	CZ	ARG	C	772	1.929	-6.773	-38.739	1.00	90.61	C
ATOM	10825	NH1	ARG	C	772	1.906	-8.079	-38.978	1.00	91.27	N
ATOM	10826	NH2	ARG	C	772	1.936	-5.935	-39.759	1.00	89.96	N
ATOM	10827	N	LEU	C	773	0.629	-4.710	-32.164	1.00	97.02	N
ATOM	10828	CA	LEU	C	773	0.423	-5.199	-30.800	1.00	97.16	C
ATOM	10829	C	LEU	C	773	-1.069	-5.199	-30.435	1.00	96.83	C
ATOM	10830	O	LEU	C	773	-1.557	-6.174	-29.874	1.00	97.78	O
ATOM	10831	CB	LEU	C	773	1.212	-4.334	-29.791	1.00	96.99	C
ATOM	10832	CG	LEU	C	773	2.759	-4.388	-29.822	1.00	96.94	C

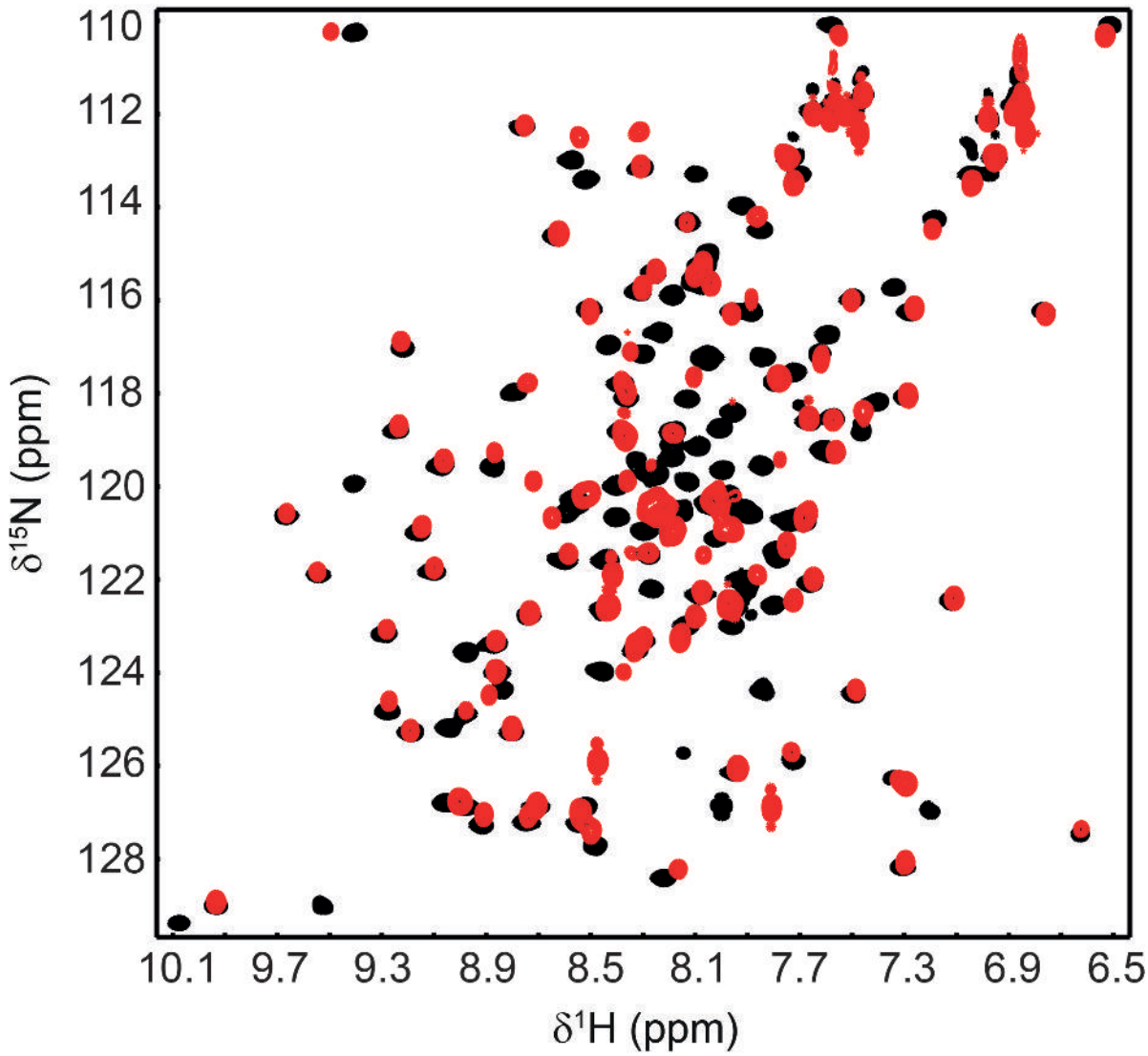
ATOM	10833	CD1	LEU	C	773	3.329	-3.401	-28.800	1.00	96.92	C
ATOM	10834	CD2	LEU	C	773	3.246	-5.800	-29.502	1.00	97.02	C
ATOM	10835	N	LEU	C	774	-1.789	-4.120	-30.755	1.00	95.96	N
ATOM	10836	CA	LEU	C	774	-3.229	-4.033	-30.473	1.00	95.46	C
ATOM	10837	C	LEU	C	774	-3.998	-5.145	-31.174	1.00	95.47	C
ATOM	10838	O	LEU	C	774	-4.830	-5.817	-30.562	1.00	96.22	O
ATOM	10839	CB	LEU	C	774	-3.806	-2.692	-30.942	1.00	93.90	C
ATOM	10840	CG	LEU	C	774	-5.346	-2.636	-30.934	1.00	92.29	C
ATOM	10841	CD1	LEU	C	774	-5.830	-2.770	-29.504	1.00	91.81	C
ATOM	10842	CD2	LEU	C	774	-5.833	-1.345	-31.545	1.00	91.22	C
ATOM	10843	N	ARG	C	775	-3.735	-5.325	-32.462	1.00	95.29	N
ATOM	10844	CA	ARG	C	775	-4.402	-6.381	-33.208	1.00	94.68	C
ATOM	10845	C	ARG	C	775	-4.209	-7.740	-32.520	1.00	94.67	C
ATOM	10846	O	ARG	C	775	-5.184	-8.404	-32.214	1.00	94.44	O
ATOM	10847	CB	ARG	C	775	-3.860	-6.445	-34.654	1.00	94.56	C
ATOM	10848	CG	ARG	C	775	-4.360	-5.341	-35.613	1.00	92.12	C
ATOM	10849	CD	ARG	C	775	-5.880	-5.356	-35.749	1.00	90.64	C
ATOM	10850	NE	ARG	C	775	-6.342	-4.568	-36.888	1.00	89.53	N
ATOM	10851	CZ	ARG	C	775	-7.616	-4.420	-37.234	1.00	88.70	C
ATOM	10852	NH1	ARG	C	775	-8.567	-5.009	-36.526	1.00	88.08	N
ATOM	10853	NH2	ARG	C	775	-7.937	-3.684	-38.292	1.00	88.26	N
ATOM	10854	N	SER	C	776	-2.950	-8.120	-32.264	1.00	94.65	N
ATOM	10855	CA	SER	C	776	-2.568	-9.411	-31.644	1.00	94.82	C
ATOM	10856	C	SER	C	776	-3.212	-9.627	-30.236	1.00	94.23	C
ATOM	10857	O	SER	C	776	-3.590	-10.761	-29.910	1.00	94.33	O
ATOM	10858	CB	SER	C	776	-1.021	-9.512	-31.521	1.00	94.77	C
ATOM	10859	OG	SER	C	776	-0.360	-9.645	-32.780	1.00	94.97	O
ATOM	10860	N	ILE	C	777	-3.334	-8.580	-29.415	1.00	93.70	N
ATOM	10861	CA	ILE	C	777	-3.901	-8.698	-28.057	1.00	92.75	C
ATOM	10862	C	ILE	C	777	-5.350	-9.252	-28.072	1.00	93.16	C
ATOM	10863	O	ILE	C	777	-5.631	-10.255	-27.404	1.00	93.55	O
ATOM	10864	CB	ILE	C	777	-3.896	-7.308	-27.305	1.00	92.10	C
ATOM	10865	CG1	ILE	C	777	-2.455	-6.821	-27.095	1.00	91.21	C
ATOM	10866	CG2	ILE	C	777	-4.534	-7.449	-25.920	1.00	91.63	C
ATOM	10867	CD1	ILE	C	777	-2.357	-5.456	-26.433	1.00	90.12	C
ATOM	10868	N	PHE	C	778	-6.256	-8.615	-28.820	1.00	93.48	N
ATOM	10869	CA	PHE	C	778	-7.673	-9.029	-28.887	1.00	93.33	C
ATOM	10870	C	PHE	C	778	-7.928	-10.029	-30.029	1.00	94.16	C
ATOM	10871	O	PHE	C	778	-8.818	-10.872	-29.934	1.00	93.84	O
ATOM	10872	CB	PHE	C	778	-8.583	-7.804	-29.105	1.00	92.40	C
ATOM	10873	CG	PHE	C	778	-8.687	-6.861	-27.918	1.00	91.76	C
ATOM	10874	CD1	PHE	C	778	-7.540	-6.454	-27.198	1.00	91.29	C
ATOM	10875	CD2	PHE	C	778	-9.943	-6.350	-27.533	1.00	91.20	C
ATOM	10876	CE1	PHE	C	778	-7.638	-5.546	-26.104	1.00	90.92	C
ATOM	10877	CE2	PHE	C	778	-10.067	-5.440	-26.441	1.00	91.24	C
ATOM	10878	CZ	PHE	C	778	-8.906	-5.037	-25.723	1.00	90.85	C
ATOM	10879	N	GLY	C	779	-7.153	-9.925	-31.108	1.00	95.41	N
ATOM	10880	CA	GLY	C	779	-7.313	-10.826	-32.242	1.00	97.01	C
ATOM	10881	C	GLY	C	779	-6.406	-10.473	-33.421	1.00	97.94	C
ATOM	10882	O	GLY	C	779	-6.809	-9.697	-34.309	1.00	98.21	O
ATOM	10883	N	GLU	C	780	-5.190	-11.038	-33.414	1.00	98.62	N
ATOM	10884	CA	GLU	C	780	-4.154	-10.817	-34.448	1.00	98.20	C
ATOM	10885	C	GLU	C	780	-4.735	-10.916	-35.859	1.00	97.77	C
ATOM	10886	O	GLU	C	780	-4.483	-10.050	-36.707	1.00	97.44	O
ATOM	10887	CB	GLU	C	780	-3.002	-11.842	-34.299	1.00	98.74	C
ATOM	10888	CG	GLU	C	780	-3.479	-13.315	-34.185	1.00	98.09	C
ATOM	10889	CD	GLU	C	780	-2.821	-14.257	-35.186	1.00	97.77	C
ATOM	10890	OE1	GLU	C	780	-3.289	-15.415	-35.292	1.00	97.28	O
ATOM	10891	OE2	GLU	C	780	-1.848	-13.846	-35.859	1.00	97.54	O
ATOM	10892	N	LYS	C	781	-5.507	-11.978	-36.099	1.00	97.33	N
ATOM	10893	CA	LYS	C	781	-6.147	-12.202	-37.392	1.00	96.31	C
ATOM	10894	C	LYS	C	781	-7.292	-11.208	-37.620	1.00	95.87	C
ATOM	10895	O	LYS	C	781	-8.485	-11.533	-37.511	1.00	96.54	O
ATOM	10896	CB	LYS	C	781	-6.675	-13.642	-37.491	1.00	95.28	C
ATOM	10897	CG	LYS	C	781	-7.331	-14.173	-36.226	1.00	93.86	C
ATOM	10898	CD	LYS	C	781	-8.176	-15.409	-36.492	1.00	93.01	C
ATOM	10899	CE	LYS	C	781	-7.389	-16.500	-37.197	1.00	92.99	C
ATOM	10900	NZ	LYS	C	781	-6.056	-16.712	-36.568	1.00	93.33	N



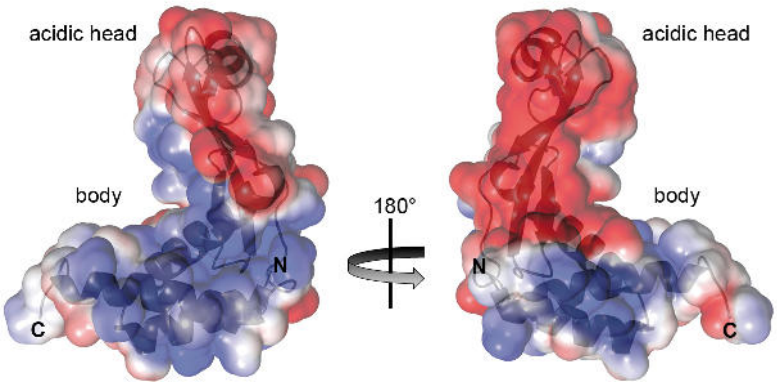
Supplementary Figure 1: Structures of NusG-NTD and RfaH-NTD. (**a**) RNAP binding site of NusG-NTD. Structure of NusG-NTD (PDB ID: 2K06) in cartoon representation, grey. Ile, Leu, and Val residues are shown as sticks with the carbon atoms of their methyl groups represented as spheres. Strongly affected methyl groups, dark red; slightly affected methyl groups, light red; unaffected methyl groups, grey; unassigned methyl groups, black. (**b**) βGL binding motif of RfaH-NTD. Structure of RfaH-NTD (PDB ID: 2OUG) in cartoon representation, orange. Residues involved in βGL binding are shown as blue sticks and labeled.



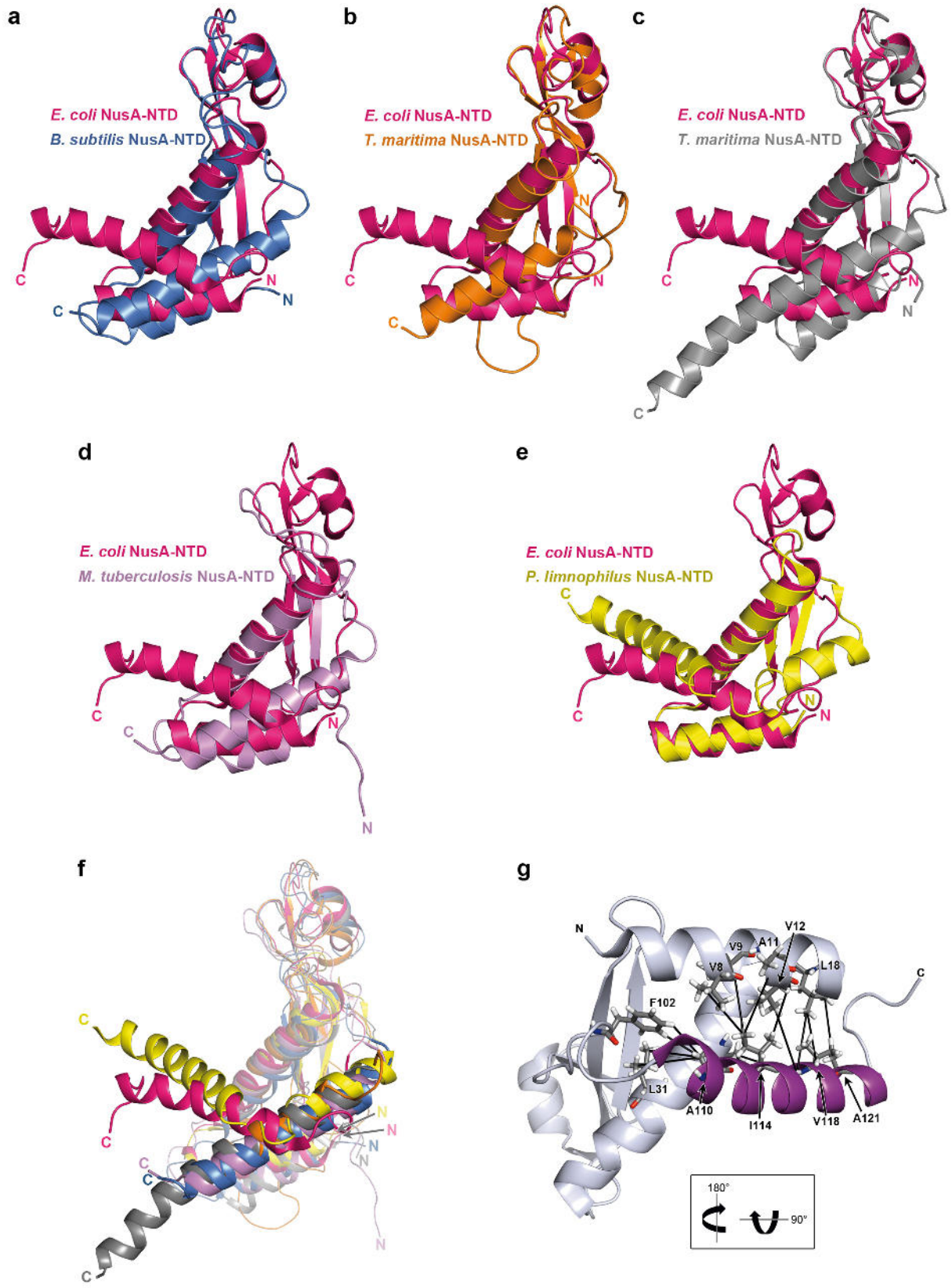
Supplementary Figure 2: Displacement of RNAP from NusE$^{\Delta}$ by NusG-CTD. 2D [^{1}H,^{15}N]-HSQC spectra of free NusB:[^{15}N]-NusE$^{\Delta}$, black, NusB:[^{15}N]-NusE$^{\Delta}$ in the presence of RNAP in equimolar concentration, red, and NusB:[^{15}N]-NusE$^{\Delta}$ in the presence of RNAP and NusG-CTD (molar ratio 1:1:1, green; 1:1:3, blue; 1:1:10, purple). Black arrows indicate the chemical shift changes that occur upon complex formation of NusG-CTD and NusB:[^{15}N]-NusE$^{\Delta 1}$.



Supplementary Figure 3: Superposition of the [^{1}H,^{15}N]-HSQC spectra of [^{15}N]-NusA-NTD(1-137), red, and [^{15}N]-NusA-NTD$^{\Delta}$, black. The protein concentration was 400 µM in each sample.



Supplementary Figure 4: Electrostatic potential molecular surface of NusA-NTD$^{\Delta}$. NusA-NTD$^{\Delta}$ in cartoon and surface representation. The electrostatic surface potential is colored from -2 kT/e, red, to +2 kT/e, blue.



Supplementary Figure 5: Comparison of NusA-NTD structures. (**a-e**) Superposition of NusA-NTD$^{\Delta}$ (pink) with (**a**) *Bs*NusA-NTD (blue, PDB ID: 2MT4, root mean square deviation

(r.m.s.d.) 1.8 Å), (**b**) *Tm*NusA-NTD (orange, PDB ID: 1HH2, r.m.s.d. 1.9 Å) (**c**) *Tm*NusA-NTD (grey, PDB ID: 1L2F, r.m.s.d. 1.7 Å), (**d**) *Mt*NusA-NTD (violet, PDB ID: 1K0R, r.m.s.d. 1.8 Å), and (**e**) *Pl*NusA-NTD (yellow, PDB ID: 4MTN, r.m.s.d. 1.4 Å). The linker helix was not used for the superpositions. **(f)** Superposition of NusA-NTD structures shown in (**a-e**). The linker helix is shown in bright colors. **(g)** NOE network fixing the position of the linker helix in NusA-NTD$^{\Delta}$ (cartoon representation, grey; the linker helix is highlighted in purple). The inset indicates how the molecule is rotated in respect to (**a**). Residues participating in the NOE network are labeled and shown as sticks (carbon atoms, dark grey; nitrogen atoms, blue; oxygen atoms, red; hydrogen atoms, white). Unambiguously identified NOEs are shown as black lines. For clarity only one NOE is displayed per methyl group (using the corresponding methyl carbon atom as center).

Supplementary References

1. Burmann, B. M. *et al*. A NusE:NusG Complex Links Transcription and Translation. *Science* **328**, 501-504 (2010).

8. Danksagung

Zunächst möchte ich mich bei Prof. Dr. Paul Rösch für die Möglichkeit bedanken meine Arbeit am Lehrstuhl Biopolymere unter hervorragenden Bedingungen durchführen zu können. Besonders bedanken möchte ich mich für das mir entgegengebrachte Vertrauen bei der Planung und Durchführung der interessanten Projekte, sowie für die Diskussionen und Anregungen, vor allem wenn ich mich in Nebensächlichkeiten verfangen hatte. Während meiner Zeit am Lehrstuhl habe ich viel gelernt.

Bei meiner Kollegin Johanna Drögemüller möchte ich mich für die außerordentlich gute Zusammenarbeit beim RNAP-Projekt bedanken. Alleine wäre die Menge an Klonierungen, Reinigungen und Messungen nicht zu bewältigen gewesen, bzw. wäre deutlich langsamer vorangegangen.

Dr. Stefan Knauer möchte ich für die Hilfe und Mitarbeit bei Postern, Vorträgen und den Manuskripten sowie für die vielen Ideen, Diskussionen und Anregungen danken. Bei Dr. Kristian Schweimer möchte ich mich für die Unterstützung bei der Vorbereitung, Auswertung und Interpretation der NMR-Spektren bedanken. Prof. Dr. Birgitta Wöhrl danke ich für ihre Hilfe bei der Planung von Klonierungen, für die Mitarbeit an den Manuskripten und für die Unterstützung bei Laborfragen. Ich danke Ulrike Persau, Ramona Heißmann und Andrea Hager für die viele Arbeit bei labororganisatorischen Dingen, für die Unterstützung bei Klonierungen und Reinigungen sowie die gute Arbeitsatmosphäre.

Rainer Hofmann danke ich für die Lösung aller Probleme rund um Computer, bzw. dafür, dass es keine größeren Probleme gab. Anja Groh möchte ich für die Unterstützung in allen Verwaltungsangelegenheiten meinen Dank aussprechen sowie für aufmunternde Worte.

Des Weiteren möchte ich meinen Freunden und Kollegen Felix Brauer, Daniel Schaal und Christian Seutter von Loetzen dafür danken, dass auch noch ein Leben außerhalb des Labors vorhanden war. Ich hoffe es hat euch mit mir auch Spaß gemacht. Anna Schneider, Maximilian Hartl, Olivia Hartl-Spiegelhauer, Philipp Weiglmeier, Berit Leo, Stephan und Sigrid Schwarzinger und Britta Zimmermann danke ich für die gute Arbeitsatmosphäre und die stetige Hilfsbereitschaft.

Zusätzlich möchte ich Prof. Max E. Gottesman (Columbia University, USA) für die Mitarbeit bei dem NusA:NusG-Manuskript danken und für den interessanten Aufenthalt in seinem Labor.
Falls ich jemanden vergessen habe, möchte ich mich auch bei ihm oder ihr bedanken und mich dafür entschuldigen. Du wirst immer einen Platz in meinem Herzen haben.

9. (Eidesstattliche) Versicherungen und Erklärungen

(§ 5 Nr. 4 PromO)

Hiermit erkläre ich, dass keine Tatsachen vorliegen, die mich nach den gesetzlichen Bestimmungen über die Führung akademischer Grade zur Führung eines Doktorgrades unwürdig erscheinen lassen.

(§ 8 S. 2 Nr. 5 PromO)

Hiermit erkläre ich mich damit einverstanden, dass die elektronische Fassung meiner Dissertation unter Wahrung meiner Urheberrechte und des Datenschutzes einer gesonderten Überprüfung hinsichtlich der eigenständigen Anfertigung der Dissertation unterzogen werden kann.

(§ 8 S. 2 Nr. 7 PromO)

Hiermit erkläre ich eidesstattlich, dass ich die Dissertation selbständig verfasst und keine anderen als die von mir angegebenen Quellen und Hilfsmittel benutzt habe. Ich habe die Dissertation nicht bereits zur Erlangung eines akademischen Grades anderweitig eingereicht und habe auch nicht bereits diese oder eine gleichartige Doktorprüfung endgültig nicht bestanden.

(§ 8 S. 2 Nr. 9 PromO)

Hiermit erkläre ich, dass ich keine Hilfe von gewerbliche Promotionsberatern bzw. -vermittlern in Anspruch genommen habe und auch künftig nicht nehmen werde.

…….

Ort, Datum, Unterschrift