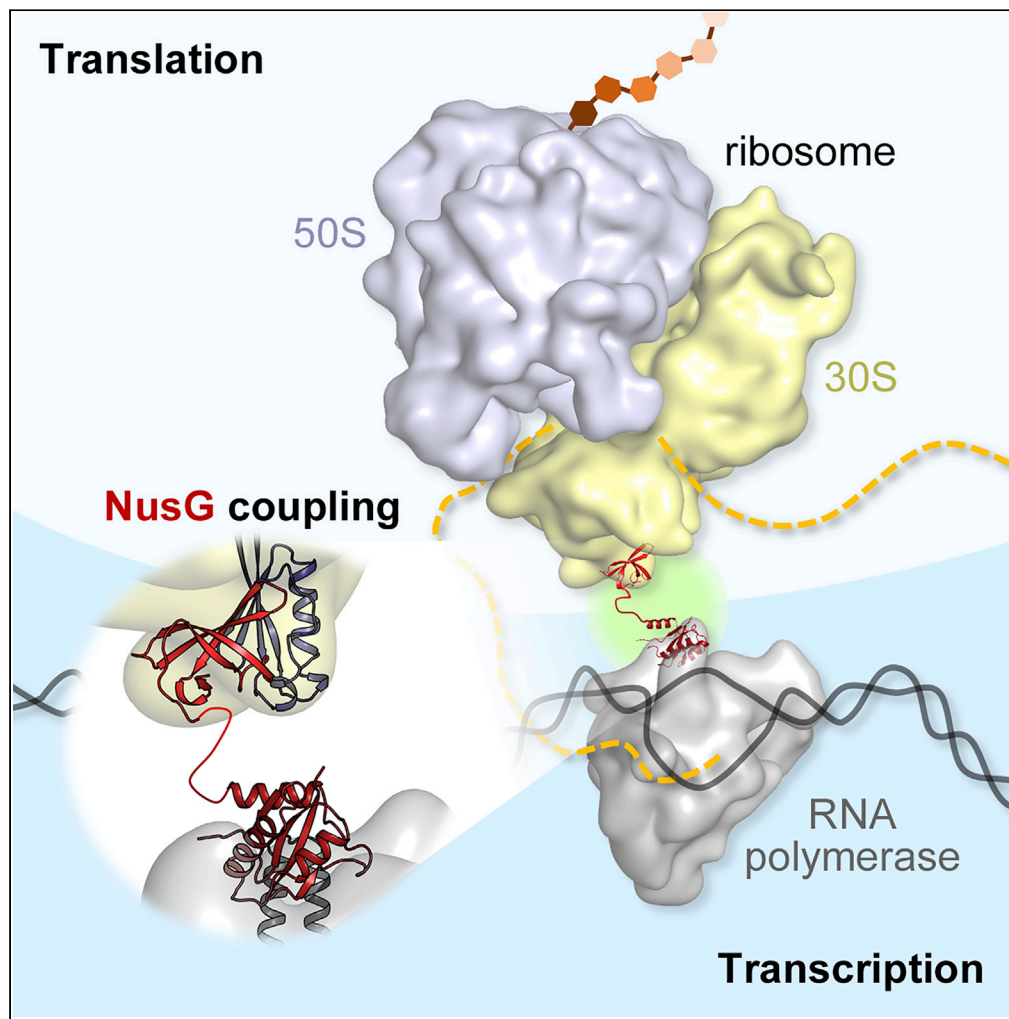


Article

Escherichia coli NusG Links the Lead Ribosome with the Transcription Elongation Complex

Robert S. Washburn,
Philipp K. Zuber,
Ming Sun, ...,
Max E. Gottesman,
Stefan H. Knauer,
Joachim Frank

meg8@cumc.columbia.edu
(M.E.G.)
stefan.knauer@uni-bayreuth.
de (S.H.K.)
jf2192@cumc.columbia.edu
(J.F.)

HIGHLIGHTS

NusG can contact RNAP
and lead ribosome
simultaneously

NusG recruitment occurs
late during transcription
and depends on
translation

NusG-mediated coupling
happens during late
translation

Washburn et al., iScience 23,
101352
August 21, 2020 © 2020 The
Authors.
[https://doi.org/10.1016/
j.isci.2020.101352](https://doi.org/10.1016/j.isci.2020.101352)

Article

Escherichia coli NusG Links the Lead Ribosome with the Transcription Elongation Complex

Robert S. Washburn,¹ Philipp K. Zuber,² Ming Sun,^{3,6} Yaser Hashem,^{4,7} Bingxin Shen,^{4,8} Wen Li,⁴ Sho Harvey,^{5,9} Francisco J. Acosta Reyes,⁴ Max E. Gottesman,^{1,4,*} Stefan H. Knauer,^{2,*} and Joachim Frank^{3,4,10,*}

SUMMARY

It has been known for more than 50 years that transcription and translation are physically coupled in bacteria, but whether or not this coupling may be mediated by the two-domain protein N-utilization substance (Nus) G in *Escherichia coli* is still heavily debated. Here, we combine integrative structural biology and functional analyses to provide conclusive evidence that NusG can physically link transcription with translation by contacting both RNA polymerase and the ribosome. We present a cryo-electron microscopy structure of a NusG:70S ribosome complex and nuclear magnetic resonance spectroscopy data revealing simultaneous binding of NusG to RNAP and the intact 70S ribosome, providing the first direct structural evidence for NusG-mediated coupling. Furthermore, *in vivo* reporter assays show that recruitment of NusG occurs late in transcription and strongly depends on translation. Thus, our data suggest that coupling occurs initially via direct RNAP:ribosome contacts and is then mediated by NusG.

INTRODUCTION

Gene expression is a universal process in all cells and consists of transcription, i.e., the synthesis of RNA based on the DNA, and—if RNA is not the final gene product—translation, i.e., the messenger RNA (mRNA)-guided synthesis of a protein. Since the late 1960s it has been known that the rates of transcription and translation are synchronized in *Escherichia coli* (*E. coli*) so that mRNA is translated while being transcribed (Das et al., 1967; Mehdi and Yudkin, 1967; Miller et al., 1970; Proshkin et al., 2010; Vogel and Jensen, 1994, 1995). This process, called transcription:translation coupling, is possible due to the lack of a physical barrier between transcription and translation in bacteria (reviewed in Conn et al., 2019). Only recently, direct physical interactions between RNA polymerase (RNAP) and the ribosome have been demonstrated (Demo et al., 2017; Fan et al., 2017; Kohler et al., 2017), consistent with earlier observations that transcriptional events may control translation activity and vice versa (Proshkin et al., 2010). As transcription and translation are closely connected to other central processes in a bacterial cell, such as DNA repair (Pani and Nudler, 2017) and protein folding (Thommen et al., 2017), transcription:translation coupling constitutes one of the key regulatory functions in bacterial gene expression.

However, there are also indications that transcription:translation coupling may involve a member of the family of N-utilization substance (Nus) G proteins, which serves as an adapter connecting RNAP and the lead ribosome (Burmam et al., 2010, 2012; Saxena et al., 2018; Zuber et al., 2019). *E. coli* NusG, member and eponym of the only universally conserved class of transcription factors (Werner, 2012), consists of two domains, an N- and a C-terminal domain (NTD and CTD), respectively, connected via a flexible linker, which move independently (Burmam et al., 2011; Mooney et al., 2009a). NusG-NTD binds RNAP and accelerates transcription elongation (Burova et al., 1995; Kang et al., 2018; Mooney et al., 2009a). Structural studies demonstrate that NusG-CTD, which is a five-stranded, antiparallel β barrel with a Kyrpides-Ouzounis-Woese motif (Kyrpides et al., 1996), is a versatile interaction platform for various transcription factors. By binding to protein S10, which is part of the 30S subunit of the ribosome, NusG may link transcription and translation (Burmam et al., 2010). Saxena et al. also demonstrated specific 1:1 binding of NusG to 70S ribosomes both *in vitro* and *in vivo* (Saxena et al., 2018). S10 is identical with transcription factor NusE and forms a ribosome-free complex with NusB, NusA, and NusG which suppresses transcription termination (Dudenhoeffer et al., 2019; Huang et al., 2019; Krupp et al., 2019; Said et al., 2017; Squires et al., 1993). Finally, NusG-CTD binds to termination factor Rho and is required for most Rho activity *in vivo* (Burmam et al., 2010; Lawson et al., 2018; Mitra et al., 2017). Transcription:translation coupling prevents Rho factor

¹Department of Microbiology & Immunology, Columbia University Medical Center, New York, NY 10032, USA

²Biochemistry IV - Biopolymers, University of Bayreuth, 95447 Bayreuth, Germany

³Department of Biological Sciences, Columbia University, New York, NY 10027, USA

⁴Department of Biochemistry and Molecular Biophysics, Columbia University Medical Center, New York, NY 10032, USA

⁵University of Michigan, Ann Arbor, MI 48109, USA

⁶Present address: University of California at San Francisco, San Francisco, CA 94158, USA

⁷Present address: NSERM U1212, Institut Européen de Chimie et Biologie, University of Bordeaux, Pessac 33607, France

⁸Present address: Bristol-Myers Squibb Pharmaceutical Co., New Brunswick, NJ 08901, USA

⁹Present address: California Institute of Technology, Pasadena, CA 91125, USA

¹⁰Lead Contact

*Correspondence: meg8@cumc.columbia.edu (M.E.G.), stefan.knauer@uni-bayreuth.de (S.H.K.), jf2192@cumc.columbia.edu (J.F.)

<https://doi.org/10.1016/j.isci.2020.101352>



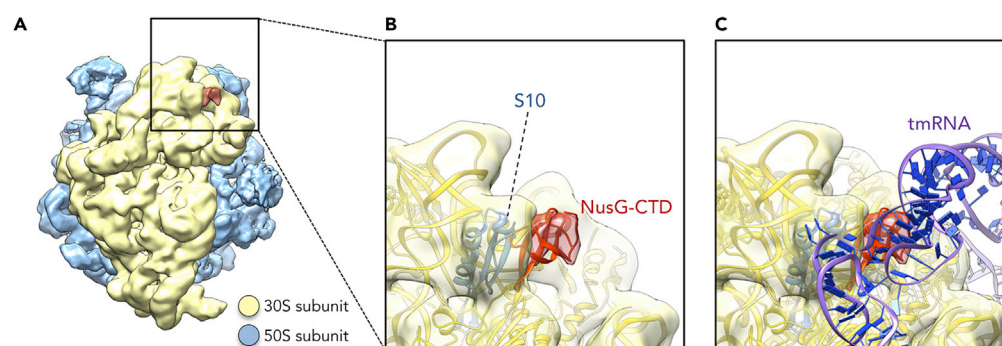


Figure 1. Structure of NusG-CTD bound to 70S Ribosome

(A) Cryo-EM density of the 70S ribosome:NusG complex (see also Table S1). The density of the 50S subunit is shown in light blue, the density of the 30S subunit in yellow, the density corresponding to NusG-CTD in red.

(B) Close-up view of the region boxed in (A). 70S (yellow), S10 (blue), and NusG-CTD (red) are in ribbon representation; cryo-EM density is shown as transparencies.

(C) Superposition of the 70S:NusG complex with the 70S:tmRNA complex (tmRNA is in ribbon representation, purple and dark blue; EMD 5234, PDB: 3IZ4). 30S and NusG-CTD are displayed as in (B).

from terminating transcription by sequestering the NusG-CTD and by blocking Rho access to RNAP via untranslated mRNA. Cryptic *E. coli* Rho-dependent terminators located within open reading frames (orfs) are revealed when ribosomes are released by polar nonsense mutations (Cardinale et al., 2008; Newton et al., 1965).

Nevertheless, there is evidence for intragenic uncoupling and Rho-dependent transcription termination in the absence of nonsense mutations: Washburn and Gottesman (2011) and Dutta et al. (2011) found that Rho resolves clashes between transcription and replication. Such conflicts are likely to occur within, rather than at the end of, genes. Uncoupling would allow Rho to release the stationary transcription elongation complexes (TECs).

Mutations in *nusE/s10* or *nusG* that uncouple transcription from translation increase sensitivity to chloramphenicol (Saxena et al., 2018). This antibiotic retards translation, breaking the bond between the lead ribosome and the TEC. Consequently, the uncoupled TEC may backtrack or terminate prematurely (Dutta et al., 2011).

In this report, we present a cryo-electron microscopy (cryo-EM) structure showing NusG binding to the S10 subunit in a 70S ribosome. The NusG-CTD binding site of S10 is also target of the ribosome-release factor, transfer-messenger RNA (tmRNA), raising the possibility that tmRNA might displace NusG at rare codons, thereby uncoupling transcription from translation (Roche and Sauer, 1999). We also show by solution-state nuclear magnetic resonance (NMR) spectroscopy that NusG, once bound to RNAP, can interact with S10 or with a complete ribosome, setting the structural basis for NusG-mediated coupling.

NusG couples transcription with translation *in vivo*, as proposed earlier (Burmann et al., 2010). Uncoupling of RNAP from the lead ribosome is enhanced when translation is compromised. Importantly, we demonstrate that uncoupled RNAP can outpace translation, leading to Rho-dependent transcription termination. This intragenic termination explains the necessity for the apparent perfect synchronization between transcription and translation (Proshkin et al., 2010).

RESULTS

Structural Evidence of NusG Binding to the Ribosomal Protein S10 on the 70S Ribosome

We assembled a NusG:70S complex by incubating 70S ribosomes with an excess of NusG and determined the structure of this complex by cryo-EM and single-particle reconstruction (Table S1). Overall, 188,127 particles were extracted from 1,327 images and ~5% of these particles showed an extra mass of density attached to the mass identified as protein S10 (Figures 1A and 1B). This additional density perfectly matches the size of NusG-CTD, suggesting that NusG binds at the site predicted from the solution

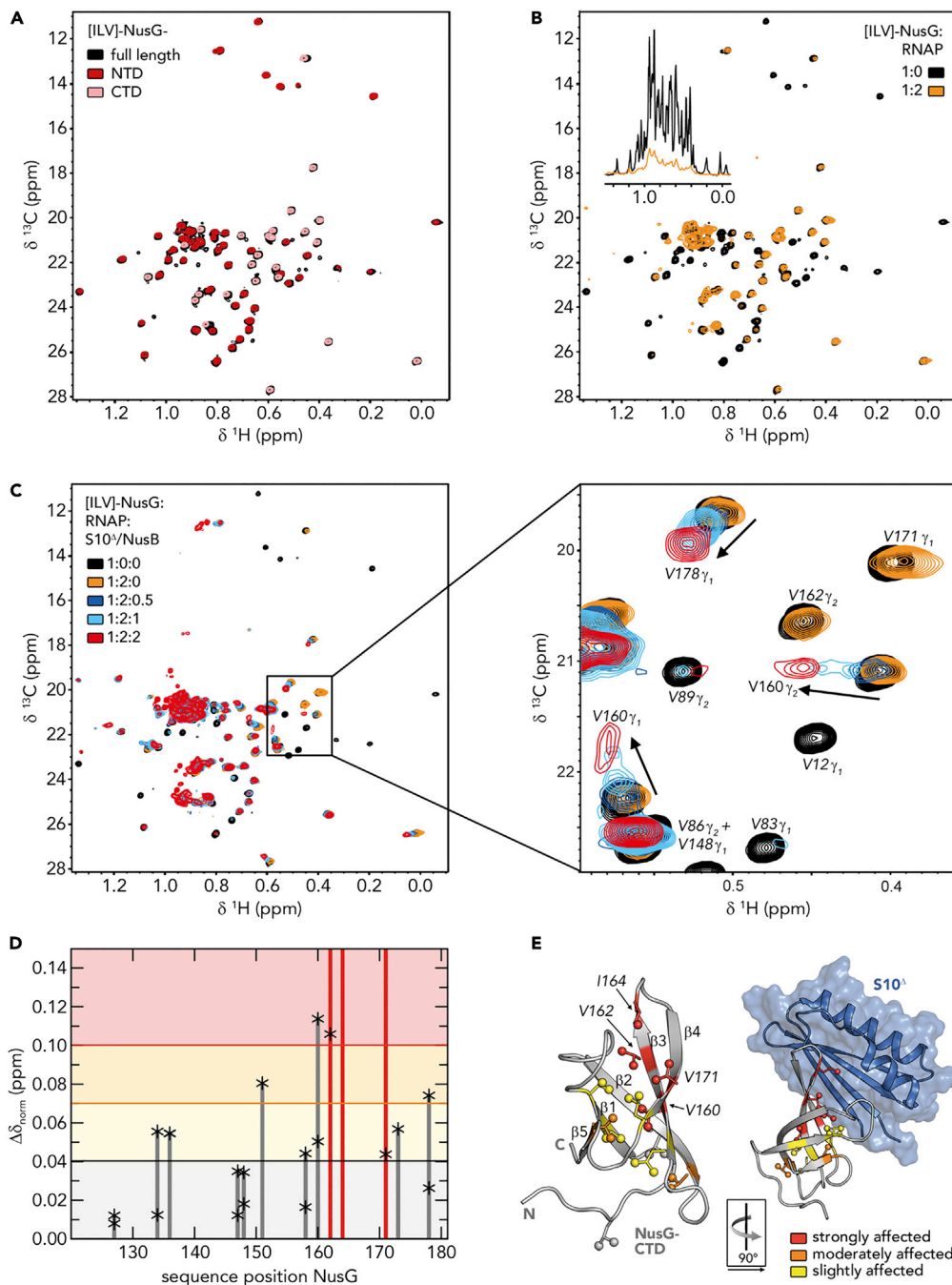


Figure 2. RNAP-Bound NusG Interacts with S10

(A) Superposition of 2D [¹H, ¹³C]-methyl-TROSY spectra of [ILV]-NusG (black, 20 μ M), [ILV]-NusG-NTD (dark red, 100 μ M), and [ILV]-NusG-CTD (light red, 30 μ M).

(B) 2D [¹H, ¹³C]-methyl-TROSY spectra of [ILV]-NusG in the absence (black, 20 μ M) and presence (orange, 18 μ M) of two equivalents of RNAP. Inset: Normalized 1D [¹H, ¹³C]-methyl TROSY spectra, colored as 2D spectra. See also Figure S1.

(C) 2D [¹H, ¹³C]-methyl-TROSY spectra of [ILV]-NusG alone (20 μ M), in the presence of a 2-fold molar excess of RNAP (18 μ M [ILV]-NusG), and upon titration of [ILV]-NusG:RNAP with 218 μ M S10^A:NusB. The molar ratio of [ILV]-NusG:RNAP:S10^A:NusB is indicated in color. The panel on the right shows an enlargement of the boxed region. Selected signals are labeled and arrows indicate chemical shift changes upon S10^A:NusB addition.

(D) [¹H, ¹³C]-methyl-TROSY-derived normalized chemical shift perturbations of [ILV]-NusG-CTD methyl group signals of RNAP-bound [ILV]-NusG upon complex formation with S10^A:NusB. Asterisks mark the values of individual methyl group

Figure 2. Continued

signals, bars represent the highest values. Red bars indicate vanishing signals. Horizontal lines are thresholds for affected methyl groups: slightly affected ($0.04 \text{ ppm} \leq \Delta\delta_{\text{norm}} < 0.07 \text{ ppm}$; black), moderately affected ($0.07 \text{ ppm} \leq \Delta\delta_{\text{norm}} < 0.1 \text{ ppm}$; orange), and strongly affected ($\Delta\delta_{\text{norm}} \geq 0.10 \text{ ppm}$; red).

(E) Mapping of affected methyl groups on the structure of isolated NusG-CTD (left; PDB ID: 2JVJ) and NusG-CTD in complex with S10^A (right; PDB ID 2KVQ). NusG-CTD is shown in ribbon (gray), S10^A in ribbon and surface (blue) representation. Affected Ile, Leu, and Val residues are colored according to (D); non-affected Ile, Leu, and Val residues are gray. Side chains of Ile, Leu, and Val residues are depicted as sticks, their methyl groups as spheres. Strongly affected Ile, Leu, and Val residues are labeled. The orientation of NusG-CTD in the complex relative to the isolated state is indicated.

NMR structure of NusG-CTD bound to the free ribosomal protein S10 in a 1:1 stoichiometry (Figures 1A and 1B [Burmam et al., 2010]). The density map reconstructed from the class of NusG:70S particles was refined to an average resolution of 6.8 Å. No density could be observed for NusG-NTD, indicating that it is flexibly bound to the NusG-CTD and does not interact with the ribosome.

During translation ribosomes may stall on incomplete mRNAs, i.e., they reach the 3' end of an mRNA without terminating, resulting in an unproductive translation complex. Together with the small protein B (SmpB) tmRNA can bind to these stalled ribosomes in order to rescue them and to tag the nascent polypeptide chain for degradation in a process called trans-translation (Weis et al., 2010). Interestingly, the NusG-CTD binding site overlaps with the region of S10 that is contacted by the tmRNA when it is bound to a ribosome in its resume state (Figure 1C [Burmam et al., 2010; Fu et al., 2010; Rae et al., 2019; Weis et al., 2010]). From this we conclude that NusG-CTD and tmRNA share binding sites on S10, raising the possibility that, in addition to releasing stalled ribosomes, tmRNA competes with NusG for ribosome binding, thus preventing NusG from maintaining a linkage between the lead ribosome and RNAP. In other words, tmRNA might be able to displace NusG and thereby facilitate uncoupled transcription.

Simultaneous Binding of NusG to S10 and RNAP

In the cryo-EM structure of *E. coli* NusG bound to a paused TEC (Kang et al., 2018) only the density of NusG-NTD was observable, indicating that NusG-CTD moves freely and does not interact with RNAP. Binding of NusG-CTD to S10 was observed both in a binary system (Burmam et al., 2010) and a λN-dependent anti-termination complex (Krupp et al., 2019; Said et al., 2017).

Since the NusG-CTD:S10 interaction is a prerequisite for NusG-mediated transcription:translation coupling, we probed this contact when NusG was bound to RNAP—but not in an antitermination context—by solution-state NMR spectroscopy. We employed NusG samples where [¹H,¹³C]-labeled methyl groups of Ile, Leu, and Val residues in perdeuterated proteins served as NMR-active probes ([ILV]-NusG) to increase sensitivity, allowing us to study large systems.

In the methyl-transverse relaxation optimized spectroscopy (methyl-TROSY) spectrum of free [ILV]-NusG (Figure 2A), signals of the NusG-NTD and NusG-CTD perfectly superimpose with the signals of the isolated [ILV]-labeled protein domains, suggesting that the domains move independently, confirming a previous report stating that there are no intramolecular domain interactions (Burmam et al., 2011). Upon addition of RNAP in a two-fold molar excess, [ILV]-NusG signals were significantly decreased in the one-dimensional methyl-TROSY spectrum (Figure 2B, inset), indicating [ILV]-NusG:RNAP complex formation. Binding of RNAP increases the molecular mass of [ILV]-NusG dramatically, resulting in enhanced relaxation, which ultimately leads to drastic line broadening and a decrease in signal intensity. Interestingly, the two-dimensional spectra revealed a non-uniform signal decrease (Figure 2B), which is caused by a combination of several effects. First, there is a general loss of signal intensity due to the increase in molecular mass upon complex formation, as discussed above. Second, upon binding, methyl groups of Ile, Leu, and Val residues located in the binding surface come into close proximity of RNAP protons. Dipole-dipole interactions contribute to relaxation processes so that the signal intensity of these methyl groups is decreased more strongly than that of methyl groups located elsewhere in [ILV]-NusG. Finally, signal intensities may be affected by chemical exchange processes. We analyzed the signal intensity of [ILV]-NusG signals in the presence of RNAP quantitatively by calculating relative signal intensities, i.e., the ratio of the remaining signal intensity of [ILV]-NusG in the presence of RNAP to the signal intensity of free [ILV]-NusG (Figure S1).

The average relative intensity of NusG-NTD signals was significantly lower than that of the linker or the NusG-CTD, suggesting that NusG-NTD binds to RNAP, whereas NusG-CTD remains flexible and moves

independently, able to interact with other partners, as indicated by the NusG:TEC structure (Kang et al., 2018). The signal intensity of all Ile, Leu, and Val residues in the RNAP binding site of NusG was completely extinguished, confirming that NusG-NTD binds to RNAP at its known binding site (Drögemüller et al., 2015; Kang et al., 2018; Krupp et al., 2019; Said et al., 2017).

To test if NusG-CTD can bind to S10 while being tethered to RNAP via NusG-NTD, we titrated the [ILV]-NusG:RNAP complex with S10^A (Figure 2C). In order to increase stability, we used an S10 variant lacking the ribosome binding loop in complex with NusB (Luo et al., 2008). Chemical shift changes of [ILV]-NusG-CTD signals upon titration of [ILV]-NusG:RNAP with S10^A:NusB were determined (Figure 2D) and affected residues were mapped onto the three-dimensional structure of NusG-CTD (Figure 2E). Strongly affected residues are located in β strands 3 and 4 as well as in the connecting loop, in agreement with the binding site observed in the binary NusG-CTD:S10^A complex (Burmenn et al., 2010). The loop between β strands 1 and 2 is also part of the NusG-CTD:S10^A binding site, but as it does not contain any Ile, Leu, or Val residues, no NMR-active probes are available in this region; nevertheless, affected residues can be found in β strand 1, directly preceding this loop. This suggests that the CTD:S10^A binding surface in the RNAP:NusG:S10^A:NusB complex is identical to the one determined in the binary system. Importantly, the NusG-NTD signals do not change when S10^A is added to the NusG:RNAP complex, indicating that S10^A binding does not release the bound RNAP.

We conclude that the S10 interaction site of NusG-CTD is accessible in the NusG:RNAP complex and thus can promote ribosome binding and formation of a ribosome:NusG:RNAP complex.

To look for a ribosome:NusG:RNAP complex, we repeated the experiment using intact 70S ribosomes instead of S10^A:NusB (Figure 3). In a first test, we titrated [ILV]-NusG with 70S ribosomes (Figure 3A). As in the [ILV]-NusG:RNAP experiment, signal intensity of [ILV]-NusG methyl groups was significantly, but not uniformly, decreased. In the presence of a 2-fold molar excess of ribosomes some NusG-NTD signals remained visible, whereas most NusG-CTD signals were nearly completely extinguished. Quantitative analysis of the [ILV]-NusG methyl group signal intensity in the presence of 0.5 equivalents of 70S ribosomes clearly shows that the relative intensity of NusG-CTD signals was in a narrow range <2%, whereas the relative intensity of NusG-NTD signals covered values from 0%–4% and was higher on average (Figure 3B). Relative intensities of zero of NusG-NTD signals can be attributed to the fact that these signals are weak even in free NusG and can thus not be quantified upon ribosome binding. Owing to the flexibility of the linker, signals corresponding to amino acids in this region had the highest relative signal intensities. From these results we conclude that NusG binds to the ribosome via its CTD, in agreement with our cryo-EM structure (Figure 1). Owing to the drastic increase in molecular mass we were unable to determine a binding site from these experiments, but nevertheless, the pattern of intensity changes of NusG-CTD signals was similar to that resulting from the titration of RNAP-bound NusG with S10^A, i.e., the most drastic decrease of signal intensity can be observed for residues 160–170, which are part of β strands 3 and 4 and the intervening loop. Consequently, we conclude that the ribosome binding site is identical with the binding site for isolated S10^A.

Next, we formed a complex of [ILV]-NusG and RNAP (molar ratio 1:2). The 2D methyl-TROSY spectrum of the complex revealed a decrease of signal intensities (Figure 3C) typical for NusG binding to RNAP (see Figure 2C), i.e., primarily NusG-CTD signals remained visible. When we then added one equivalent of 70S ribosomes nearly all [ILV]-NusG signals were diminished (e.g., the signal corresponding to I164, which is in the loop responsible for ribosome binding; Figure 3C). Strikingly, the spectrum differs from the spectrum of [ILV]-NusG in the presence of 70S ribosome (Figure 3A). These results can be explained by three scenarios: (1) NusG-NTD is bound to RNAP, NusG-CTD is bound to a ribosome, and the ribosome directly interacts with RNAP; (2) NusG-NTD is bound to RNAP, NusG-CTD is bound to the ribosome, but the ribosome does not interact with RNAP; (3) NusG-NTD is bound to RNAP, the ribosome directly interacts with RNAP, and NusG-CTD is free but is in the vicinity of the ribosome. To exclude the last scenario we repeated the experiment using a NusG variant, NusG^{F165A}, in which F165, essential for ribosome binding (Burmenn et al., 2010; Knowlton et al., 2003), is substituted by an Ala. Having ensured that the amino acid substitution does not influence the structure of NusG (Figure S2A) we tested in a control experiment [ILV]-NusG^{F165A} binding to S10^A. Indeed, we detected no interaction (Figures S2B and S2C). When we added 70S ribosomes to a preformed [ILV]-NusG^{F165A}:RNAP complex (molar ratio 1:2), the spectrum corresponding to the [ILV]-NusG^{F165A}:RNAP complex did not change significantly and, in particular, NusG-CTD signals

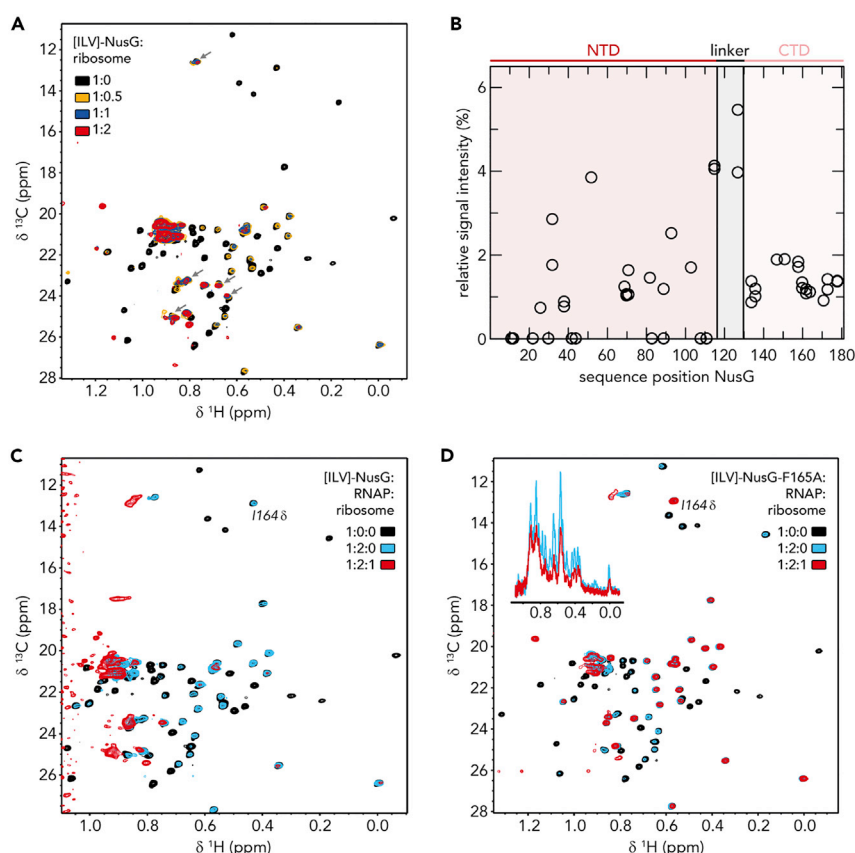


Figure 3. RNAP-bound NusG Interacts with the 70S Ribosome

(A and B) NusG interacts with 70S ribosome via its CTD. (A) 2D ^1H , ^{13}C -methyl-TROSY spectra of free [ILV]-NusG (11 μM , black) and [ILV]-NusG in the presence of 70S ribosome (molar ratio [ILV]-NusG:ribosome = 1:0.5 (6.6 μM [ILV]-NusG, orange); = 1:1 (7.5 μM [ILV]-NusG, blue); = 1:2 (4 μM [ILV]-NusG, red)). Arrows indicate [ILV]-NusG-NTD signals that are well visible in the [ILV]-NusG:ribosome complex. (B) Quantitative analysis of [ILV]-NusG methyl group signal intensities in the presence of 0.5 equivalents of 70S ribosome. Relative signal intensities are plotted versus the sequence position of NusG. The domain organization of NusG is indicated above the diagram.

(C) 2D ^1H , ^{13}C -methyl-TROSY spectra of [ILV]-NusG (11 μM , black), [ILV]-NusG in the presence of RNAP (molar ratio 1:2, 6 μM [ILV]-NusG, blue), and [ILV]-NusG in the presence of RNAP and 70S ribosome (molar ratio 1:2:1, 6 μM [ILV]-NusG, red). (D) 2D ^1H , ^{13}C -methyl-TROSY spectra of [ILV]-NusG^{F165A} (20 μM , black), [ILV]-NusG^{F165A} in the presence of RNAP (molar ratio 1:2, 6 μM [ILV]-NusG^{F165A}, blue), and [ILV]-NusG^{F165A} in the presence of RNAP and 70S ribosome (molar ratio 1:2:1, 6 μM [ILV]-NusG^{F165A}, red). The inset shows the normalized 1D spectra of the corresponding titration step.

See also Figure S2.

remained visible, suggesting that the ribosome was not bound (Figure 3D). However, the general decrease in signal intensity indicates a direct RNAP:ribosome interaction. Thus, we conclude that NusG can serve as physical linker between ribosome and RNAP, although it remains elusive if a direct interaction between RNAP and a ribosome occurs in this NusG-coupled complex.

Translation Promotes NusG Attachment to TEC

Chromatin immunoprecipitation (ChIP) analysis showed that NusG binds to TEC well after transcription and translation initiation (Mooney et al., 2009b). Thus, we asked whether translation was, in fact, required for attachment of NusG to the TEC. To approach this question, we examined the effects of translation on NusG-mediated Rho-dependent termination within the *lac* operon (Figure 4A, Table 1) as NusG recruitment to the TEC is necessary for efficient Rho-dependent termination. Rho-dependent termination occurs within *lacZ* both *in vitro* (Burns and Richardson, 1995) and, upon the introduction of *lacZ* nonsense mutations, *in vivo* (Adhya and Gottesman, 1978; Newton et al., 1965). Polarity was measured using a probe to

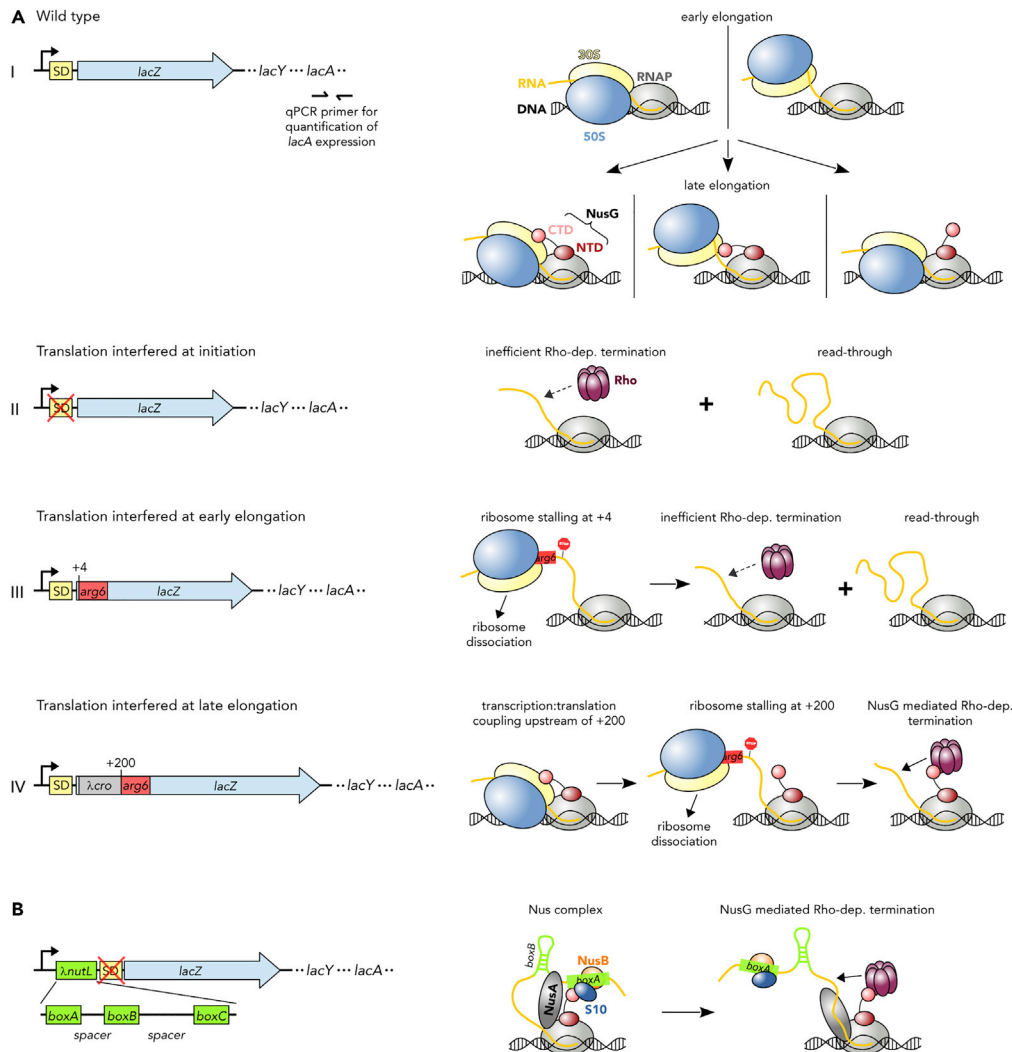


Figure 4. Translation Is Required for NusG Recruitment to the TEC

(A and B) Left: Organization of the *E. coli* *lac* operon in strains MDS42 (A-I; wild type *lacZ*), RSW1225 (A-II; mutant [inactive] *lacZ* SD sequence), RSW1245 (A-III; in-frame insertion of six rare Arg codons [*arg6*] at position +4 of *lacZ*), RSW1276 (A-IV; in-frame insertion of λ cro and six rare Arg codons at position +4 of *lacZ* [equivalent to *arg6* being at position +200 of the gene]), and RSW1297 (B; λ nutL site upstream of mutant *lacZ* SD sequence). *lacY* and *lacA* are only indicated for clarity. qPCR primers specific to the 3' end of *lacA* (position indicated in A-I) were used to measure mRNA levels and thereby readthrough of *lacA* (see Table 1). Right: Schemes of possible effects on transcription:translation coupling and Rho-dependent termination within *lacZ*. A-I, top: Ribosomes are recruited in the early elongation phase, leading to a directly coupled RNAP:ribosome complex (left) or uncoupled transcription and translation (right). A-I, bottom: NusG is recruited in late elongation, resulting in a NusG-coupled complex with (left) or without (middle) direct RNAP:ribosome contacts, or modifying the pre-existing RNAP:ribosome complex without establishing an CTD:S10 interaction (right). A-II: Failure of NusG recruitment results in inefficient Rho-dependent termination and high *lacZ* readthrough. A-III: *arg6* stops the translating ribosome at position +4, whereas transcription elongation proceeds (left), resulting in ribosome dissociation and no NusG recruitment. Transcription proceeds and is only inefficiently terminated by Rho (right). A-IV: NusG couples transcription and translation (left) until *arg6* stops the ribosome at position +200 (middle), allowing efficient, NusG-stimulated Rho-dependent termination (right). (B) λ nutL recruits NusA, NusG, and the S10/NusB dimer, creating a Nus complex. NusG can thus support Rho-dependent termination.

lacA, comparing mRNA levels with or without treatment with the Rho inhibitor bicyclomycin (BCM). Wild-type (WT) cells revealed no detectable termination (Table 1 and Figure 4A-I), which may be attributed to (1) sequestering of NusG-CTD by the ribosome, (2) binding of the ribosome to the nascent RNA, or (3) both. In all scenarios, however, the presence of the translating ribosome prevents Rho binding. We interfered with

Strain	<i>lacZ</i>	<i>nutL</i>	Fold Increase of RNA Level (BCM ⁻)	Fold Increase of RNA Level (BCM ⁺)	RT (%)
MDS42	wt	-	.25 ± 0.04	.26 ± 0.03	96 ± 19
RSW1225	SD ⁻	-	.12 ± 0.03	.56 ± 0.10	21 ± 7
RSW1245	arg(6)—early	-	.13 ± 0.01	.49 ± 0.02	27 ± 2
RSW1276	arg(6)—late	-	<.001 ± 0.002	.12 ± 0.003	<1
RSW1297	SD ⁻	+	.01 ± 0.006	.59 ± 0.01	2 ± 1

Table 1. NusG Couples Late after Transcription Initiation

Expression of *lacZ* was induced for 20 min from the *lac* operon with 1 mM IPTG. Where indicated, Rho-dependent termination was inhibited by adding 100 μg/mL BCM 1 min prior to induction. Readthrough was calculated from the fold increase of *lacA* RNA compared with *ompA* RNA in the presence or absence of BCM. RNA levels were measured using qRT-PCR and the fold increase was calculated using the $\Delta\Delta C_t$ method (Livak and Schmittgen, 2001). Values are the average of ≥ 3 independent experiments, each carried out in duplicate. RSW1225 carries two G to A mutations in the *lacZ* ribosome-binding site. RSW1245 carries an insertion of six rare arginine codons (atg-acc-atg-AGG-AGA-CGA-AGG-AGA-CGA) at the amino terminus of *lacZ*. RSW1276 contains six rare arginine codons 200 nt distal to the start site of translation. RSW1297 carries an insertion of λ *nutL* immediately 5' to the mutated ribosome binding site.

translation initiation by mutating the ribosome-binding site, i.e., the Shine-Dalgarno (SD) sequence (Figure 4A-II), or translation elongation by introducing six successive rare arginine codons at two different locations in *lacZ* (Figure 4A-III and IV). Introduction of two G to A mutations in the *lacZ* SD sequence prohibits translation initiation of *lacZ* (Figure 4A-II). *lacA* mRNA measurements gave a readthrough of 21%, indicating that Rho-dependent termination occurs, but was inefficient in the absence of translation of *lacZ* mRNA. Introduction of the six in-frame rare arginine residues at the +4 position of *lacZ* (Figure 4A-III and Table 1) allowed 27% readthrough, i.e., Rho-dependent termination is present but still inefficient if translation of *lacZ* mRNA is interfered with at early elongation. In contrast, introduction of the rare arginine residues 200 nucleotides (nt) from the start site of transcription (Figure 4A-IV and Table 1) resulted in high polarity, yielding <1% readthrough. As efficient Rho-dependent termination requires NusG our results suggest that NusG binding to TEC occurs late and is dependent on translation.

To confirm the hypothesis that NusG failed to attach to TEC in the absence of translation, we asked if a complex comprising Nus factors A, B, and E (Nus complex) assembled at a λ *nutL* site was able to recruit NusG so that it associates with TEC. Accordingly, we introduced the λ *nutL* site just upstream of the flawed *lacZ* SD sequence and measured *lacA* mRNA level (Figure 4B and Table 1). Indeed, Rho-dependent termination was highly efficient, indicating that NusG had been recruited to TEC. Thus, counterintuitively, the Nus complex, which normally suppresses transcription termination in ribosomal (*rrn*) operons (Dudenhoeffer et al., 2019; Huang et al., 2019; Squires et al., 1993) and, together with λ N, on the phage λ chromosome, stimulates termination in this case.

We finally demonstrated that reduced termination efficiency in the mutant with the non-functional SD sequence was due to the failure of NusG recruitment to the TEC. In this assay we monitored Rho-dependent termination in a fusion construct that carries λ *cro*, the λ *nutR* site, the Rho-dependent λ *tR1* terminator, and a *lacZ* reporter, with *lacZ* expression being heat-inducible (Figure 5). Termination at the λ *tR1* site is poor when *cro* is translated, as seen with the *cro ms27* fusion (Table 2 and Figure 5A-I. In the presence of an intact SD sequence we used *cro ms27*, where codon 27 carries a missense mutation so that the resulting protein is non-functional. The 3' end of *cro* is adjacent to the λ *tR1* terminator, limiting the amount of free RNA available for Rho attachment if *cro* mRNA is translated. When λ *cro* carried an SD mutation translation initiation was ablated, but nevertheless there was significant termination at λ *tR1* (Table 2 and Figure 5A-II). Formation of the Nus complex at λ *nutR* allows NusG recruitment and efficient termination. In the absence of NusB, the complex does not assemble and there is extensive readthrough at λ *tR1*.

The *boxA69* mutation also reduces Nus complex formation at λ *nutR* and, like the *nusB⁻* mutation, enhances readthrough of λ *tR1* (Table 3 and Figure 5B). In this experiment, we suppressed termination at λ *tR1* with λ N antitermination factor instead of BCM. Finally, we showed that expression of *nusG-NTD*, which competes with NusG for binding to RNAP, enhances readthrough (Table 4 and Figure 5C). Taken together,

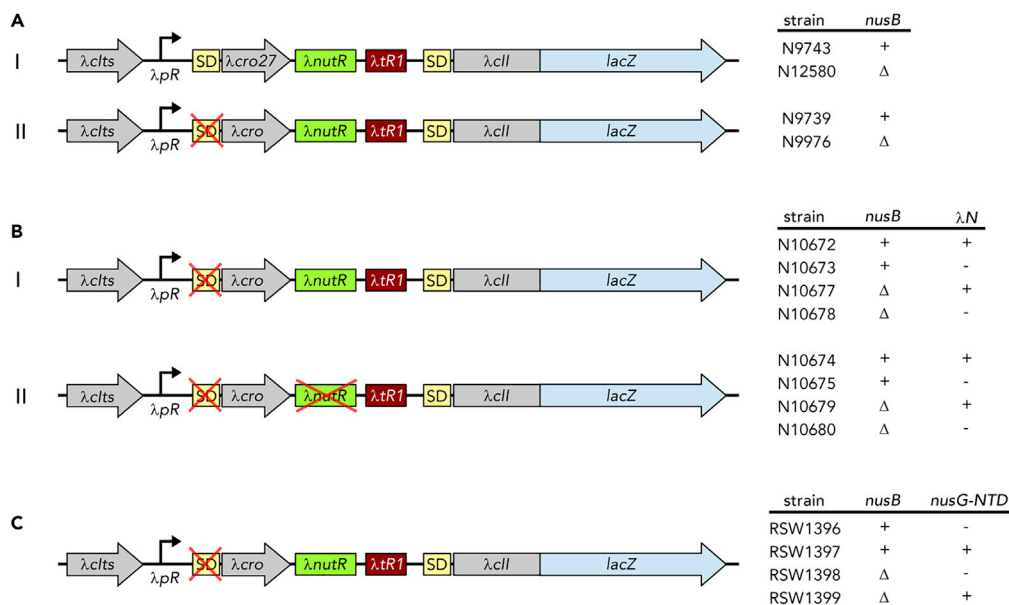


Figure 5. NusG Can Be Recruited via a Nus Complex

Genetic constructs used to monitor NusG-mediated Rho-dependent termination are shown with the corresponding strains and their properties indicated on the right side. Transcription is started from the λpR promoter, followed by WT- λcro or λcro carrying a missense mutation at codon 27 ($\lambda cro27$), a WT or mutant $\lambda nutR$ site, the Rho-dependent terminator $\lambda tR1$, and a $\lambda cII::lacZ$ transcriptional fusion with a corresponding SD site. All strains encode a temperature-sensitive λcI construct ($\lambda cIts$) to allow temperature-controlled induction of gene expression from the λpR promoter. (A) Nus complex formation compensates the lack of an SD sequence. (B) BoxA mutation impairs NusG recruitment (C) Uncoupling by NusG-NTD. λN^+ strains listed in (B) further encode the λN protein; in (B-II) the non-functional $\lambda nutR$ sequence was generated by the *boxA69* mutation; NusG-NTD for strains listed in (C) was supplied from plasmid pRM442. See also Tables 2, 3, and 4.

these results strongly support the idea that NusG can be supplied by the Nus complex assembled at $\lambda nutR$ in the absence of translation, inducing Rho-dependent termination at $\lambda tR1$.

DISCUSSION

Structural Basis of NusG-Mediated Transcription: Translation

We determined a cryo-EM structure of a NusG:70S complex showing binding of one molecule NusG per ribosome, consistent with previous results (Saxena et al., 2018). NusG binds to the S10 protein on the 30S subunit via its CTD as indicated by the study of isolated NusG-CTD and S10^A (Burmam et al., 2010); density for NusG-NTD was not observable, suggesting that it remains flexible. We must attribute the low occupancy of the NusG-CTD on the 70S ribosome in the cryo-EM experiment to weak binding adversely affected by the conditions of sample preparation. Notably, although tmRNA contacts the ribosome at various sites, the binding of NusG-CTD and tmRNA on S10 seems to be mutually exclusive. This suggests a model in which uncoupling at rare codons, at which tmRNA releases ribosomes, is promoted by tmRNA-induced release of NusG (Roche and Sauer, 1999). The freed NusG:TEC complex exposes the NusG-CTD and is then subject to Rho-dependent transcription termination. This model, however, requires that the affinity of tmRNA for S10 is higher than for the NusG-CTD:S10 interaction. This could be the subject of further studies. Alternatively, tmRNA binding to S10 might only occur once NusG-mediated coupling has been disrupted owing to ribosome stalling, allowing transcription to continue while tmRNA rescues the ribosome.

Simultaneous binding of NusG to S10^A and RNAP has been demonstrated by solution-state NMR studies, confirming the S10^A binding site on NusG-CTD as identified in a binary NusG-CTD:S10^A system (Figure 2) (Burmam et al., 2010). Moreover, we show that NusG can bind isolated RNAP and isolated 70S ribosome concurrently. Although this is not an actively transcribing and translating system, our data provide the first

Strain	<i>cro</i>	<i>nusB</i>	β -Galactosidase Activity (BCM ⁻) in Miller Units	β -Galactosidase Activity (BCM ⁺) in Miller Units	RT (%)
9,743	<i>ms27</i>	+	530 \pm 3	680 \pm 5	78 \pm 0.7
12,580	<i>ms27</i>	Δ	890 \pm 11	1,150 \pm 15	77 \pm 1.4
9,739	SD ⁻	+	141 \pm 3	613 \pm 25	23 \pm 1.1
9,976	SD ⁻	Δ	1,191 \pm 17	1,290 \pm 36	92 \pm 3.0

Table 2. NusG Coupling at *nutR* Requires NusB

Expression of *lacZ* was induced from a chromosomal *cll::lacZ* transcriptional fusion (λ clts-*pR-cro-nutR-trl1-cll::lacZ*) by incubating at 42°C for 30 min. N9743 and N12580 carry a missense mutation at *cro* codon 27; N9739 and 9,976 have a G to C mutation in the *cro* SD sequence (SD⁻); N12580 and N9976 are deleted for *nusB*. Where indicated, BCM was added to 100 μ g/mL prior to induction of *lacZ* expression. Readthrough (RT) was calculated from the ratio of β -galactosidase activity (in Miller units) in the presence or absence of BCM (BCM⁺ and BCM⁻, respectively). Miller units from ≥ 3 independent experiments were averaged.

direct structural evidence consistent with NusG-mediated transcription:translation coupling. The flexibility of the linker between the NusG-NTD and the NusG-CTD permits these interactions.

The operon-specific *E. coli* NusG paralog, RfaH, likewise simultaneously binds S10^A and RNAP in the context of a paused TEC (Burmam et al., 2012; Zuber et al., 2019). RfaH, which also comprises an NTD and a flexibly connected CTD (Belogurov et al., 2007; Burmam et al., 2012), uses the same binding sites as NusG to interact with RNAP and S10 (Burmam et al., 2010, 2012; Kang et al., 2018; Sevostyanova et al., 2011; Zuber et al., 2019). However, RfaH, unlike NusG, complexes with TEC early after transcription initiation, when TEC pauses at an operon polarity suppressor (*ops*) site, a representative of the *E. coli* consensus pause sequence (Larson et al., 2014; Vvedenskaya et al., 2014). Located in the untranslated leader region of RfaH-controlled operons, *ops* is responsible for sequence-specific recruitment of RfaH (Zuber et al., 2018). Importantly, RfaH-dependent operons lack a consensus SD sequence. To initiate translation, RfaH recruits a ribosome to these mRNAs, making coupling essential for translation activation and efficient gene expression (Burmam et al., 2012). The binding modes of RfaH and NusG to RNAP and S10 are very similar, indicating that coupling as observed for RfaH can also be mediated by NusG and vice versa. However, once recruited, RfaH excludes NusG (Kang et al., 2018), thus preventing intra-operon Rho-dependent transcription termination in RfaH-controlled operons (see Artsimovitch and Knauer, 2019).

Recruitment of NusG Requires Translation and Stimulates Rho-Dependent Termination

We have confirmed the results of Mooney et al. that NusG binds to TEC only after significant RNA synthesis (Mooney et al., 2009b). As postulated by these authors, binding depends on active translation of the mRNA. Thus, efficient Rho-dependent transcription termination, which requires the attachment of Rho to the NusG-CTD, does not occur at the end of an untranslated gene. We have shown that the failure of NusG to bind TEC is responsible for the absence of termination. Thus, placing a λ *nut* site at the start of the gene recruits NusG and restores termination. At present, it is not understood why NusG appears to be delivered to TEC by ribosomes *in vivo*, whereas it binds directly to RNAP in a purified system lacking ribosomes. A possible explanation would be that NusG attaches to RNAP discontinuously in an on-and-off mode in the untranslated leader region and that the NusG:RNAP interaction is only stabilized when the ribosome is coupled upon translation initiation. We should recall that NusG has two binding sites in the coupled system, which significantly increases its avidity.

Multiple Modes of Transcription:Translation Coupling

A direct connection between transcription and translation was first predicted in 1964 (Byrne et al., 1964), and later it was shown that transcription:translation coupling is necessary to coordinate gene expression as well as to maintain genome stability (McGary and Nudler, 2013). In 1970, Miller et al. performed electron microscopy analyses of lysed *E. coli* cells (Miller et al., 1970). They demonstrated that all mRNA molecules are connected to the *E. coli* genome and that the ribosome at the newly synthesized end of a polyribosome is almost always immediately adjacent to the putative RNAP molecule. Finally, they concluded that translation is completely coupled with transcription. Coupling could allow RNAP to monitor the translation rate

Strains	<i>boxA</i>	<i>nusB</i>	β -Galactosidase Activity (λN^-) in Miller Units	β -Galactosidase Activity (λN^+) in Miller Units	RT (%)
10,673; 10,672	+	+	125 \pm 1	946 \pm 23	13 \pm 0.3
10,675; 10,674	69	+	1,212 \pm 30	2,211 \pm 87	55 \pm 2.5
10,678; 10,677	+	Δ	2,874 \pm 24	2,616 \pm 103	100 \pm 4.4
10,680; 10,679	69	Δ	1,896 \pm 25	2,416 \pm 80	78 \pm 2.8

Table 3. *BoxA* Mutations Block NusG Coupling at *nutR*

Expression of *lacZ* was induced from a chromosomal *cll::lacZ* transcriptional fusion ($\lambda clts$ -*pR-cro* (*SD*) -*nutR*-*tR1-cll::lacZ*) by incubation at 42°C for 30 min. Strains N10672, N10674, N10677, and N10679 express λN , which encodes the transcription termination inhibitor λN . *boxA*69 and Δ *nusB* strain numbers are indicated in Table 3. RT was calculated from the ratio of β -galactosidase activity in the presence or absence of λN (λN^+ and λN^- , respectively). Miller units from ≥ 3 independent experiments were averaged.

while providing newly synthesized mRNA to the ribosome. The structural basis of this coupling is, however, still only poorly understood. Our results strongly suggest that NusG may mediate coupling ("indirect coupling"). Since NusG attaches to the TEC downstream to the translation initiation site, the coupled transcription:translation complex must initially consist of a ribosome bound directly to TEC. This "direct coupling" mode is in agreement with both structural and biochemical data (Demo et al., 2017; Fan et al., 2017; Kohler et al., 2017). A cryo-EM structure of a directly coupled complex has been published where a translating ribosome collided into a stalled RNAP, forming a so-called expressome (Figure 6A [Kohler et al., 2017]). In this complex, RNAP directly binds to the 30S subunit with the RNA exit region of RNAP docking onto the ribosome near the mRNA tunnel entry between ribosomal proteins S3, S4, and S5, allowing the mRNA exiting from RNAP to enter directly into the ribosome. Another cryo-EM structure showed an RNAP:30S complex generated by mixing 30S subunit with a 3-fold excess of RNAP (Figure 6B [Demo et al., 2017]). In this structure RNAP is bound to the 30S subunit near the mRNA binding site between the head and the platform domains, contacting ribosomal proteins S1, S2, S18, S21, and hairpin loop 40 of 16S rRNA, in agreement with cross-linking data (Fan et al., 2017). Strikingly, this position is located more than 80 Å from the binding site observed in the expressome structure, i.e., on the opposite side of the 30S head. Importantly, it ensures that RNAP interacts with the cytosolic side of the 30S ribosomal subunit so that the nascent RNA exiting from RNAP is directly guided to the entry site on the ribosome. Assuming that both the RNAP:30S complex and the expressome correspond to active coupling complexes, the structures indicate that multiple coupling modes exist, which involve massive relocalization of RNAP relative to ribosome.

Interestingly, neither the RNAP:30S nor the expressome structures allow NusG- (or RfaH-) mediated coupling: the linker of NusG/RfaH is too short (Figure 6). However, as the cryo-EM structures suggest that the position of RNAP on the 30S subunit might be flexible, these structures could be snapshots of distinct situations during translation. Thus, we suggest that, at some distance downstream of the translation initiation site, NusG recognizes and enters the coupled complex, rearranging its structure.

While our manuscript was under review, two preprints have been published reporting several structures of coupled complexes from *E. coli* (Wang et al., 2020; Webster et al., 2020). Overall, the structures indicate that there are indeed various types of transcription:translation coupling modes, both direct and indirect coupling. The coupling mode depends on the length of the mRNA separating the RNAP active site and the ribosomal P-site and is determined by the position of RNAP relative to the ribosome. For example, the structure of the collided complex (see above [Kohler et al., 2017]) was confirmed, but this coupling mode may be relevant only under certain conditions or when the RNA spacer is very short. In NusG-coupled complexes (Wang et al., 2020; Webster et al., 2020) NusG bridges RNAP and ribosome with NusG-NTD contacting the RNAP at the expected binding site (Kang et al., 2018) and NusG-CTD interacting with S10. This is similar to what was found for the binary NusG-CTD:S10 system (Burmam et al., 2012) and is in agreement with our data. As compared with the collided complex, RNAP is significantly rotated relative to the ribosome, but, interestingly, no stable contacts between RNAP and ribosome were observable (Webster et al., 2020). However, more coupling modes are possible, emphasizing the complexity of the interplay between transcription and translation (Wang et al., 2020; Webster et al., 2020).

Strain	<i>nusG-NTD</i>	<i>nusB</i>	β -Galactosidase Activity (BCM ⁻) in Miller Units	β -Galactosidase Activity (BCM ⁺) in Miller Units	RT (%)
RSW1396	–	+	247 ± 5	862 ± 2	29 ± 0.6
RSW1397	+	+	944 ± 3	1,013 ± 7	93 ± 0.7
RSW1398	–	Δ	2,013 ± 33	2,314 ± 55	87 ± 2.5
RSW1399	+	Δ	2,360 ± 37	2,760 ± 150	86 ± 4.8

Table 4. NusG-NTD Uncouples Transcription and Translation

Expression of *lacZ* was induced from a chromosomal *cll::lacZ* transcriptional fusion (λ clts-*pR-cro*(SD⁻)-*nutR-tR1-cll::lacZ*) by incubating at 42°C for 30 min. *nusG-NTD* expression was induced from the plasmid pRM442 *tac* promoter with 1 mM IPTG for 10 min prior to induction of *lacZ* in strains RSW1397 and RSW 1399. Strains RSW1396 and RSW1398 carried an empty vector (*ptrc99A*) and were exposed to IPTG as above. Where indicated BCM was added to 100 μ g/mL prior to induction of *lacZ*. RT was calculated from the ratio of β -galactosidase activity in the presence or absence of BCM (BCM⁺ and BCM⁻, respectively). Miller units from ≥ 3 independent experiments were averaged.

Conclusion

In summary, multiple transcription:translation coupling modes exist. Based on our results, we hypothesize that direct coupling between the ribosome and TEC occurs during translation initiation and early elongation, whereas NusG-mediated coupling is established later in translation in *E. coli*. Notably, the transition between the different coupling modes requires significant rearrangement of the relative position of RNAP and ribosome, with NusG serving as additional anchor point to restrict the freedom of RNAP movement relative to the ribosome.

Coupling may synchronize transcription and translation; in particular, the leading ribosome may “push” RNAP to overcome transcriptional pauses while RNAP could “pull” the ribosome to prevent/escape translational pausing (Wang et al., 2020). In NusG-coupled complexes the small transcription factor may serve as cushion, conferring the system the flexibility necessary to keep transcription and translation synchronized, even if these processes are regulated differently or occur at different rates.

Limitations of the Study

- (1) RNA extraction efficiencies are subject to some variability.
- (2) Cryo-EM modeling: (a) The accuracy of the atomic model is limited by the cryo-EM resolution and the low occupancy of NusG on the ribosome. With the current resolution of the map, it is impossible to identify any rearrangement of side chains involved in interactions that stabilize the formation of the NusG:70S complex. The limited resolution also has the consequence that we had to rely on the accuracy of published structures used in the fitting and docking. (b) Furthermore, we modeled the 30S subunit structure only, based on a published atomic structure, even though the 70S-NusG complex was visualized. However, there are only minor differences between the structures of the isolated and 50S-bound 30S subunit, and the geometry of NusG binding determined by the modeling should not be affected.
- (3) NMR: (a) We used isolated RNAP and isolated ribosome in our experiments, and, consequently, not an actively transcribing/translating system. Thus, we provide only structural evidence that NusG can link RNAP and ribosome in the absence of nucleic acids. The situation might, theoretically, be different *in vivo* (although this is not very probable). (Already indicated in the section “Structural basis of NusG-mediated transcription:translation.”) (b) Although we find that NusG can serve as linker between RNAP and ribosome we cannot distinguish between two scenarios: (i) NusG links RNAP and ribosome without direct interactions between RNAP and ribosome and (ii) NusG links RNAP and ribosome, but RNAP and ribosome make direct contacts in this coupled complex (already indicated at the end of section “Simultaneous binding of NusG to S10 and RNAP”).

Resource Availability

Lead Contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Joachim Frank (jf2192@cumc.columbia.edu).

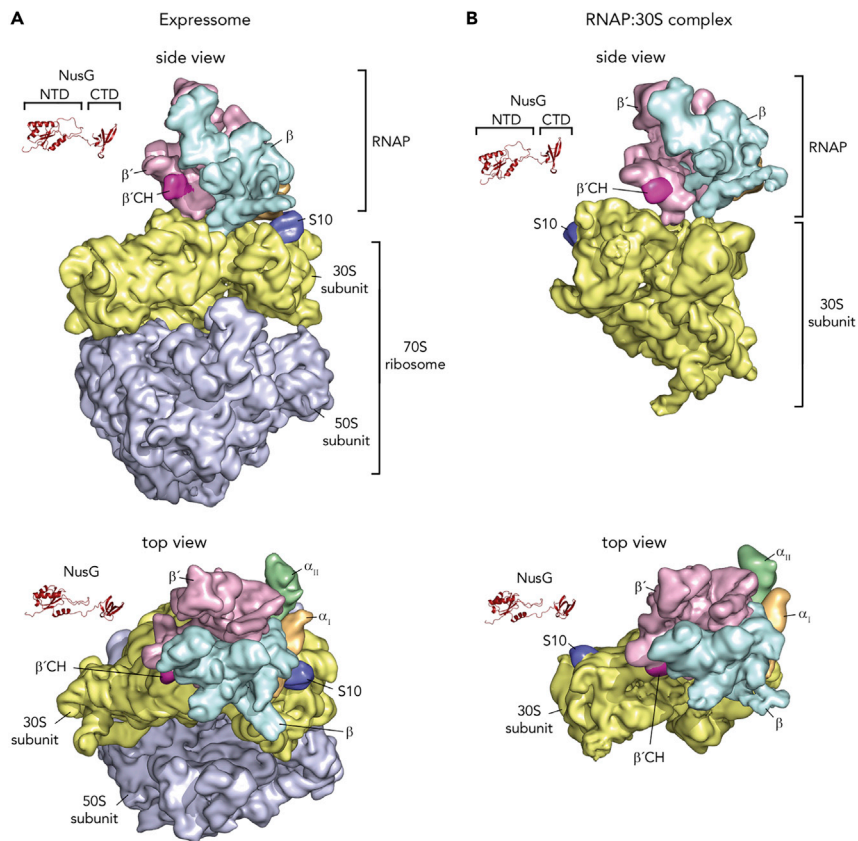


Figure 6. Structures of Coupled Complexes

Structures of the expressome (A) and an RNAP:30S complex (B) determined by cryo-EM. RNAP and ribosomal subunits are in surface representation, NusG is shown as ribbon. α_I , orange; α_{II} green; β , cyan; β' , light violet; 30S, yellow; 50S, light blue; β' CH, pink; S10, dark blue. PDB IDs: expressome, 5MY1 and 6O9J; RNAP:30S, 6AWD; NusG-NTD: 2K06; NusG-CTD: 2JVV.

Materials Availability

This study did not generate new unique reagents.

Data and Code Availability

For the 70S:NusG complex visualized by cryo-EM, only the region of 30S:NusG-CTD was modeled for simplicity. Electron densities and the final atomic model for the 70S:NusG complex have been deposited in the Electron Microscopy Data Bank (EMDB) and the Protein Data Bank (PDB) with accession codes EMD-22143 and 6XE0.

METHODS

All methods can be found in the accompanying [Transparent Methods supplemental file](#).

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at <https://doi.org/10.1016/j.isci.2020.101352>.

ACKNOWLEDGMENTS

We gratefully acknowledge the help of D. Shapoval, M. Bubunencko, N. Costantino, and D. Court, and we are deeply thankful for numerous useful discussions with P. Röscher. Supported by HHMI and NIH R01 GM29169 (to J.F.), NIH R01 GM037219 (to M.E.G.), and the German Research Foundation grant Ro617/21-1 (to P. Röscher). S.H. was supported by an Amgen Fellowship.

AUTHOR CONTRIBUTIONS

B.S., M.S., P.K.Z., R.S.W., S.H., and Y.H. performed the experiments and processed the data. J.F., M.E.G., M.S., P.K.Z., R.S.W., S.H.K., and Y.H. designed the experiments and interpreted the results. F.J.A.R., M.S., and W.L. modeled the atomic structure of 30S:NusG-CTD. J.F., M.E.G., M.S., P.K.Z., and S.H.K. wrote the paper. This publication was funded by the German Research Foundation and the University of Bayreuth in the funding program “Open Access Publishing”.

DECLARATION OF INTERESTS

The authors declare no competing financial interests.

Received: March 11, 2020

Revised: June 4, 2020

Accepted: July 3, 2020

Published: August 21, 2020

REFERENCES

- Adhya, S., and Gottesman, M. (1978). Control of transcription termination. *Annu. Rev. Biochem.* 47, 967–996.
- Artsimovitch, I., and Knauer, S.H. (2019). Ancient transcription factors in the news. *MBio* 10, e01547–18.
- Belogurov, G.A., Vassilyeva, M.N., Svetlov, V., Klyuyev, S., Grishin, N.V., Vassilyev, D.G., and Artsimovitch, I. (2007). Structural basis for converting a general transcription factor into an operon-specific virulence regulator. *Mol. Cell* 26, 117–129.
- Burmann, B.M., Schweimer, K., Luo, X., Wahl, M.C., Stitt, B.L., Gottesman, M.E., and Rösch, P. (2010). A NusE:NusG complex links transcription and translation. *Science* 328, 501–504.
- Burmann, B.M., Scheckenhofer, U., Schweimer, K., and Rösch, P. (2011). Domain interactions of the transcription-translation coupling factor *Escherichia coli* NusG are intermolecular and transient. *Biochem. J.* 435, 783–789.
- Burmann, B.M., Knauer, S.H., Sevostyanova, A., Schweimer, K., Mooney, R.A., Landick, R., Artsimovitch, I., and Rösch, P. (2012). An α helix to β barrel domain switch transforms the transcription factor RfaH into a translation factor. *Cell* 150, 291–303.
- Burns, C.M., and Richardson, J.P. (1995). NusG is required to overcome a kinetic limitation to Rho function at an intragenic terminator. *Proc. Natl. Acad. Sci. U S A* 92, 4738–4742.
- Burova, E., Hung, S.C., Sagitov, V., Stitt, B.L., and Gottesman, M.E. (1995). *Escherichia coli* NusG protein stimulates transcription elongation rates *in vivo* and *in vitro*. *J. Bacteriol.* 177, 1388–1392.
- Byrne, R., Levin, J.G., Bladen, H.A., and Nirenberg, M.W. (1964). The *in vitro* formation of a DNA-ribosome complex. *Proc. Natl. Acad. Sci. U S A* 52, 140–148.
- 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.
- Conn, A.B., Diggs, S., Tam, T.K., and Blaha, G.M. (2019). Two old dogs, one new trick: a review of RNA polymerase and ribosome interactions during transcription-translation coupling. *Int. J. Mol. Sci.* 20, 2595.
- Das, H.K., Goldstein, A., and Lowney, L.I. (1967). Attachment of ribosomes to nascent messenger RNA in *Escherichia coli*. *J. Mol. Biol.* 24, 231–245.
- Demo, G., Rasouly, A., Vasilyev, N., Svetlov, V., Loveland, A.B., Diaz-Avalos, R., Grigorieff, N., Nudler, E., and Korostelev, A.A. (2017). Structure of RNA polymerase bound to ribosomal 30S subunit. *Elife* 6, e28560.
- 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.
- Dudenhoeffer, B.R., Schneider, H., Schweimer, K., and Knauer, S.H. (2019). SuhB is an integral part of the ribosomal antitermination complex and interacts with NusA. *Nucleic Acids Res.* 47, 6504–6518.
- Dutta, D., Shatalin, K., Epshtein, V., Gottesman, M.E., and Nudler, E. (2011). Linking RNA polymerase backtracking to genome instability in *E. coli*. *Cell* 146, 533–543.
- Fan, H., Conn, A.B., Williams, P.B., Diggs, S., Hahn, J., Gamper, H.B., Hou, Y.-M., O’Leary, S.E., Wang, Y., and Blaha, G.M. (2017). Transcription-translation coupling: direct interactions of RNA polymerase with ribosomes and ribosomal subunits. *Nucleic Acids Res.* 45, 11043–11055.
- Fu, J., Hashem, Y., Wower, I., Lei, J., Liao, H.Y., Zwiebe, C., Wower, J., and Frank, J. (2010). Visualizing the transfer-messenger RNA as the ribosome resumes translation. *EMBO J.* 29, 3819–3825.
- Huang, Y.-H., Said, N., Loll, B., and Wahl, M.C. (2019). Structural basis for the function of SuhB as a transcription factor in ribosomal RNA synthesis. *Nucleic Acids Res.* 47, 6488–6503.
- Kang, J.Y., Mooney, R.A., Nedialkov, Y., Saba, J., Mishanina, T.V., Artsimovitch, I., Landick, R., and Darst, S.A. (2018). Structural basis for transcript elongation control by NusG family universal regulators. *Cell* 173, 1650–1662.e14.
- Knowlton, J.R., Bubunenko, M., Andrykovitch, M., Guo, W., Routzahn, K.M., Waugh, D.S., Court, D.L., and 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.
- Kohler, R., Mooney, R.A., Mills, D.J., Landick, R., and Cramer, P. (2017). Architecture of a transcribing-translating expressome. *Science* 356, 194–197.
- Krupp, F., Said, N., Huang, Y.-H., Loll, B., Bürger, J., Mielke, T., Spahn, C.M.T., and Wahl, M.C. (2019). Structural basis for the action of an all-purpose transcription Anti-termination factor. *Mol. Cell* 74, 143–157.e5.
- Kyrpides, N.C., Woese, C.R., and Ouzounis, C.A. (1996). KOW: a novel motif linking a bacterial transcription factor with ribosomal proteins. *Trends Biochem. Sci.* 21, 425–426.
- Larson, M.H., Mooney, R.A., Peters, J.M., Windgassen, T., Nayak, D., Gross, C.A., Block, S.M., Greenleaf, W.J., Landick, R., and Weissman, J.S. (2014). A pause sequence enriched at translation start sites drives transcription dynamics *in vivo*. *Science* 344, 1042–1047.
- Lawson, M.R., Ma, W., Bellecourt, M.J., Artsimovitch, I., Martin, A., Landick, R., Schulten, K., and Berger, J.M. (2018). Mechanism for the regulated control of transcription by a universal adapter protein. *Mol. Cell* 71, 1–12.
- Livak, K.J., and Schmittgen, T.D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2⁻(Delta Delta C(T)) Method. *Methods* 25, 402–408.
- Luo, X., Hsiao, H.-H., Bubunenko, M., Weber, G., Court, D.L., Gottesman, M.E., Urlaub, H., and Wahl, M.C. (2008). Structural and functional analysis of the *E. coli* NusB-S10 transcription antitermination complex. *Mol. Cell* 32, 791–802.
- McGary, K., and Nudler, E. (2013). RNA polymerase and the ribosome: the close relationship. *Curr. Opin. Microbiol.* 16, 112–117.

- Mehdi, Q., and Yudkin, M.D. (1967). Coupling of transcription to translation in the induced synthesis of beta-galactosidase. *Biochim. Biophys. Acta* 149, 288–290.
- Miller, O.L., Hamkalo, B.A., and Thomas, C.A. (1970). Visualization of bacterial genes in action. *Science* 169, 392–395.
- Mitra, P., Ghosh, G., Hafeezunnisa, M., and Sen, R. (2017). Rho protein: roles and mechanisms. *Annu. Rev. Microbiol.* 71, 687–709.
- Mooney, R.A., Schweimer, K., Rösch, P., Gottesman, M., and Landick, R. (2009a). 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.
- Mooney, R.A., Davis, S.E., Peters, J.M., Rowland, J.L., Ansari, A.Z., and Landick, R. (2009b). Regulator trafficking on bacterial transcription units *in vivo*. *Mol. Cell* 33, 97–108.
- Newton, W.A., Beckwith, J.R., Zipser, D., and Brenner, S. (1965). Nonsense mutants and polarity in the lac operon of *Escherichia coli*. *J. Mol. Biol.* 14, 290–296.
- Pani, B., and Nudler, E. (2017). Mechanistic insights into transcription coupled DNA repair. *DNA Repair (Amst.)* 56, 42–50.
- Proshkin, S., Rahmouni, A.R., Mironov, A., and Nudler, E. (2010). Cooperation between translating ribosomes and RNA polymerase in transcription elongation. *Science* 328, 504–508.
- Rae, C.D., Gordiyenko, Y., and Ramakrishnan, V. (2019). How a circularized tmRNA moves through the ribosome. *Science* 363, 740–744.
- Roche, E.D., and Sauer, R.T. (1999). SsrA-mediated peptide tagging caused by rare codons and tRNA scarcity. *EMBO J.* 18, 4579–4589.
- Said, N., Krupp, F., Anedchenko, E., Santos, K.F., Dybkov, O., Huang, Y.-H., Lee, C.-T., Loll, B., Behrmann, E., Bürger, J., et al. (2017). Structural basis for λ N-dependent processive transcription antitermination. *Nat. Microbiol.* 2, 17062.
- Saxena, S., Myka, K.K., Washburn, R., Costantino, N., Court, D.L., and Gottesman, M.E. (2018). *Escherichia coli* transcription factor NusG binds to 70S ribosomes. *Mol. Microbiol.* 108, 495–504.
- 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.
- Squires, C.L., Greenblatt, J., Li, J., Condon, C., and Squires, C.L. (1993). Ribosomal RNA antitermination *in vitro*: requirement for Nus factors and one or more unidentified cellular components. *Proc. Natl. Acad. Sci. U S A* 90, 970–974.
- Thommen, M., Holtkamp, W., and Rodnina, M.V. (2017). Co-translational protein folding: progress and methods. *Curr. Opin. Struct. Biol.* 42, 83–89.
- Vogel, U., and Jensen, K.F. (1994). Effects of guanosine 3',5'-bis(diphosphate) (ppGpp) on rate of transcription elongation in isoleucine-starved *Escherichia coli*. *J. Biol. Chem.* 269, 16236–16241.
- Vogel, U., and Jensen, K.F. (1995). Effects of the antiterminator BoxA on transcription elongation kinetics and ppGpp inhibition of transcription elongation in *Escherichia coli*. *J. Biol. Chem.* 270, 18335–18340.
- Vvedenskaya, I.O., Vahedian-Movahed, H., Bird, J.G., Knoblauch, J.G., Goldman, S.R., Zhang, Y., Ebright, R.H., and Nickels, B.E. (2014). Interactions between RNA polymerase and the “core recognition element” counteract pausing. *Science* 344, 1285–1289.
- Wang, C., Molodtsov, V., Firlar, E., Kaelber, J., Blaha, G., Su, M., and Ebright, R.H. (2020). Structural basis of transcription-translation coupling. *bioRxiv*. <https://doi.org/10.1101/2020.03.01.972380>.
- Washburn, R.S., and Gottesman, M.E. (2011). Transcription termination maintains chromosome integrity. *Proc. Natl. Acad. Sci. U S A* 108, 792–797.
- Webster, M.W., Takacs, M., Zhu, C., Vidmar, V., Eduljee, A., Abdelkareem, M., and Weixlbaumer, A. (2020). Structural basis of transcription-translation coupling and collision in bacteria. *bioRxiv*. <https://doi.org/10.1101/2020.03.01.971028>.
- Weis, F., Bron, P., Giudice, E., Rolland, J.-P., Thomas, D., Felden, B., and Gillet, R. (2010). tmRNA-SmpB: a journey to the centre of the bacterial ribosome. *EMBO J.* 29, 3810–3818.
- 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.
- Zuber, P.K., Artsimovitch, I., NandyMazumdar, M., Liu, Z., Nedialkov, Y., Schweimer, K., Rösch, P., and Knauer, S.H. (2018). The universally-conserved transcription factor RfaH is recruited to a hairpin structure of the non-template DNA strand. *Elife* 7, e36349.
- Zuber, P.K., Schweimer, K., Rösch, P., Artsimovitch, I., and Knauer, S.H. (2019). Reversible fold-switching controls the functional cycle of the antitermination factor RfaH. *Nat. Commun.* 10, 702.

Supplemental Information

***Escherichia coli* NusG Links the Lead Ribosome with the Transcription Elongation Complex**

Robert S. Washburn, Philipp K. Zuber, Ming Sun, Yaser Hashem, Bingxin Shen, Wen Li, Sho Harvey, Francisco J. Acosta Reyes, Max E. Gottesman, Stefan H. Knauer, and Joachim Frank

Supplemental Tables

Table S1: Electron Microscopy and Modeling. Related to Figure 1.

Data collection	Polara-F30
Micrographs	1,327
Picked Particles	188,127
Voltage (KV)	300
Pixel size (Å/pixel)	1.66
Structure	
Component	70S:NusG
Particles	17,122
FSC 0.143 (Å)	6.8
Model Refinement	
Component	30S:NusG
Program/Protocol	MDFF/Phenix
Used in refinement (Å)	6.8
Average B-factors (Å ²)	556.01
Avg B-fac Prot/RNA (Å ²)	571.78/546.93
R.m.s deviations:	
Bonds (Å)	0.009
Angles (deg)	1.123
Validation	
Molprobity score	1.61
Clashcore, all atoms	2.59
Rotamer outliers (%)	1.85
Ramachandran plot:	
Outliers (%)	0.17
Allowed (%)	5.15
Favored (%)	94.68
Composition	
Non hydrogen atoms	52,035
Protein residues	2,413
RNA bases	1,539
Ligands	0
Accession codes	
EMDB	EMD-22143
PDB	6XE0

Supplemental Figures

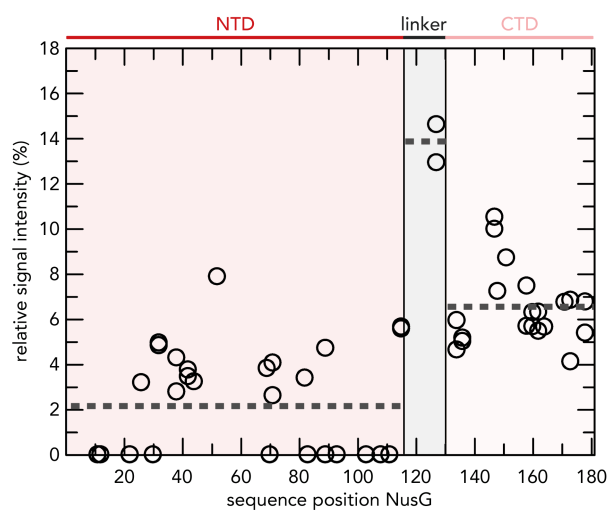


Figure S1: Binding of [ILV]-NusG to RNAP. Related to Figure 2B. $[^1\text{H}, ^{13}\text{C}]$ -methyl-TROSY derived relative signal intensities of [ILV]-NusG methyl groups after addition of two equivalents of RNAP (see Fig. 2B). Dashed horizontal lines indicate average relative signal intensities of NusG-NTD, the linker, and NusG-CTD (domain organization is indicated at the top).

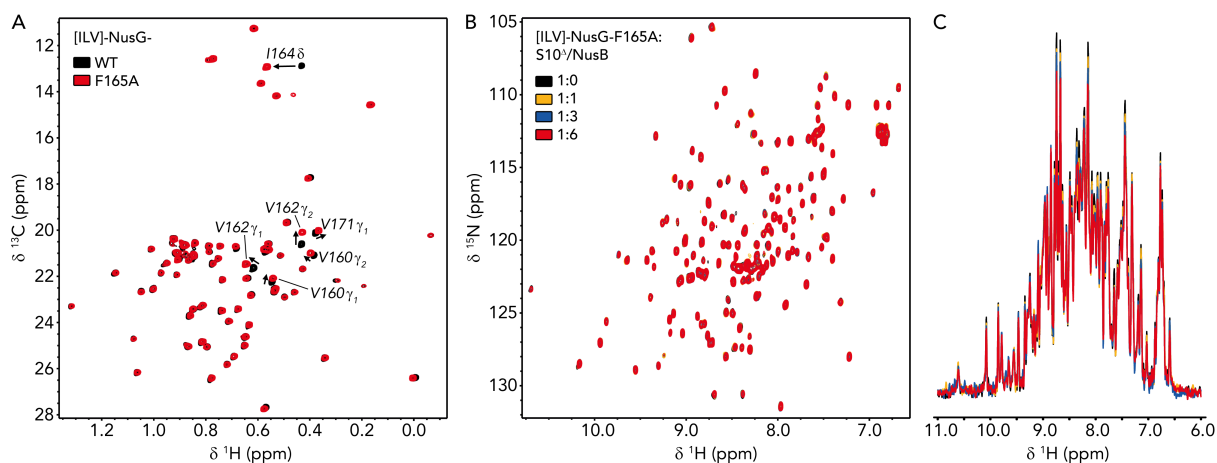


Figure S2: NusG^{F165A} does not interact with S10::NusB. Related to Figure 3. **(A)** 2D [^1H , ^{13}C]-methyl-TROSY spectra of [ILV]-NusG (11 μM , black) and [ILV]-NusG^{F165A} (20 μM , red). Arrows and labels indicate NusG-CTD methyl groups affected in their resonance frequencies by the F165A amino acid substitution. **(B,C)** 2D **(B)** and normalized 1D **(C)** [^1H , ^{15}N]-HSQC spectra of 20 μM [ILV]-NusG^{F165A} upon titration with 432 μM S10::NusB (colors as indicated).

Transparent Methods

Strain construction. Standard bacteriological techniques used in strain construction (e.g., transformation, transduction and media preparation) are as described in (Silhavy et al., 1984). Standard molecular biology techniques were as described in Sambrook and Russell (Sambrook and Russel, 2001). N10780 was constructed by P1 transduction of *rpoC-his:kanR nusGF165A* from NB885 into MDS42. N11158 was constructed by P1 transduction of Δ *ssrA::camR* from RSW943 into MDS42. N11816 was constructed by P1 transduction of Δ *relA::kanR* from RLG847 into N11158. RSW1008 was constructed by P1 transduction of Δ *ssrA::camR* from RSW943 into N4837. RSW1010 was constructed by P1 transduction of *rpoC-his:kanR nusGF165A* from NB885 into N4837. RSW1012 was constructed by P1 transduction of Δ *ssrA::camR* from RSW943 into RSW1010. RSW1175 was constructed by P1 transduction of Δ *relA::kanR* and Δ *spoT::camR* from RLG847 into MDS42. RSW1245 was generated using recombineering (Sharan et al., 2009) to introduce six rare arginine codons (atg-acc-atg-AGG-AGA-CGA-AGG-AGA-CGA-att-acg-gat) into the 5' end of *lacZ* in MDS42 changing the amino acid sequence of the aminotermminus from MTMITD to MTMRRRRRRITD with six inefficiently translated arginine codons. RSW1225 was produced using recombineering to introduce two G to A mutations in the ribosome binding site of *lacZ* in MDS42. This resulted in a change from ...TTCACACAGGAAACAGCTatgaccatg... to ...TTCACACACC AAACAGCTatgaccatg... inactivating the ribosome binding site. RSW1225 is *lac*⁻.

Cloning. The plasmid encoding NusG^{F165A} (pET11a_*nusG-F165A*) was generated by site-directed mutagenesis according to the QuikChange Site-Directed Mutagenesis Kit protocol (Stratagene), using vector pET11a_*nusG* (Burmann et al., 2011) as template and primers Fw_NusG-F165A (5' GTG TCT GTT TCT ATC GCG GGT CGT GCG ACC CCG 3') and

Rv_NusG-F165A (5' CGG GGT CGC ACG ACC CGC GAT AGA AAC AGA CAC 3'; both primers were obtained from metabion, Martinsried, Germany).

Protein production and isotopic labeling. For the production of unlabeled proteins, bacteria were grown in lysogeny broth (LB) medium. [^1H , ^{13}C]-labeling of methyl groups of Ile, Leu, and Val residues in perdeuterated proteins was accomplished by growing bacteria in minimal medium M9 (Meyer and Schlegel, 1983; Sambrook and Russel, 2001) prepared with increasing amounts of D_2O (0 % (v/v), 50 % (v/v), 100 % (v/v); Eurisotop, Saint-Aubin, France) and $(^{15}\text{NH}_4)_2\text{SO}_4$ (CortecNet, Voisins-Le-Bretonneux, France) and d_7 -glucose (Cambridge Isotope Laboratories, Inc., Tewksbury, USA) as sole nitrogen and carbon sources, respectively. Amino acid precursors (60 mg/l 2-keto-3- d_3 -4- ^{13}C -butyrate and 100 mg/l 2-keto-3-methyl- d_3 -3- d_1 -4- ^{13}C -butyrate; Eurisotop, Saint-Aubin, France) were added 1 h prior to induction. Expression and purification protocols were identical to those of non-labeled proteins.

Production of full-length NusG and NusG^{F165A} for NMR studies was based on (Burmann et al., 2011). For expression, *E. coli* BL21 (λ DE3) cells (Novagen, Madison, WI, USA) harboring plasmids pET11a_nusG or pET11a_nusG-F165A (encoding tag-less *E. coli* NusG or NusG^{F165A}, respectively) were grown in medium containing ampicillin (100 $\mu\text{g/ml}$) to an optical density at 600 nm (OD_{600}) of 0.8 at 37 °C. Overexpression was subsequently induced by addition of 1 mM isopropyl-1-thio- β -D-galactopyranoside (IPTG) and continued for 4 h at 37 °C. Cells were harvested by centrifugation (6,000 x g), resuspended in buffer A^{NusG} (50 mM tris(hydroxymethyl)aminomethane (Tris)/HCl (pH 7.5), 250 mM NaCl) supplemented with ½ protease inhibitor cocktail tablet (cOmplete, EDTA-free, Roche Diagnostics, Mannheim, Germany) and a small amount of DNase I (AppliChem GmbH, Darmstadt, Germany), and lysed using a microfluidizer. After clearing the lysate by centrifugation (13,000 x g, 30 min, 4 °C), streptomycin sulfate was added to a final

concentration of 1 % (w/v) and the solution was stirred for 30 min at room temperature to precipitate nucleic acids. Following centrifugation, the supernatant was successively supplemented with $(\text{NH}_4)_2\text{SO}_4$ at 4 °C under continuous stirring to a final concentration of 60 % (w/v). The pellet was collected by centrifugation, dissolved in buffer B^{NusG} (10 mM Tris/HCl (pH 7.5)) and subsequently dialyzed against the same buffer overnight. The solution was then applied to a 5 ml HeparinFF column (GE Healthcare, Munich, Germany) equilibrated with buffer B^{NusG}. After washing with 20 column volumes (CVs) buffer B^{NusG}, proteins were eluted using a step gradient from 50 mM – 1 M NaCl in buffer B^{NusG}. Target protein containing fractions were combined, concentrated by ultrafiltration (molecular weight cut-off (MWCO): 5 kDa), and then applied to a HiLoad S75 size exclusion chromatography (SEC) column (GE Healthcare, Munich, Germany) equilibrated with buffer C^{NusG} (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES; pH 7.5), 100 mM NaCl). Fractions containing pure NusG/NusG^{F165A} were combined, concentrated by ultrafiltration (MWCO: 5 kDa), flash-frozen in liquid nitrogen and stored at -80 °C.

NusG for cryoEM was produced based on (Saxena et al., 2018). In brief, *E. coli* BL21 (λ DE3) cells (Novagen, Madison, WI, USA) harboring plasmid pRM431, which codes for NusG fused to a hexahistidine tag at its C-terminus, were grown in LB medium containing 100 $\mu\text{g/ml}$ ampicillin at 37 °C. Upon reaching an OD_{600} of 0.5 *nusG-his₆* expression was induced by IPTG addition (0.5 mM) and cells were harvested 3 hours later by centrifugation (4,347 x g, 15 min, 4 °C). The cell pellet was resuspended in buffer D^{NusG} (50 mM Tris/HCl (pH 7.5), 150 mM NH_4Cl), $\frac{1}{2}$ a tablet protease inhibitor (EDTA-free, Sigma-Aldrich) was added, and cells were disrupted by four freeze-thaw cycles in a dry-ice ethanol bath and water at room temperature followed by sonication (4 x 15 s pulses). The lysate was centrifuged (12,000 x g, 30 min, 4 °C) and the crude extract was added to Ni-NTA resin (GE Healthcare) equilibrated with buffer D^{NusG} (5 ml resin / 10 ml crude extract). After overnight incubation with rotation at 4 °C the resin was packed in 2 ml columns and each column was washed

with 5 CVs buffer D^{NusG}. Elution was carried out with increasing imidazole concentration (100-500 mM imidazole in buffer D^{NusG}). Target protein containing fractions were combined, dialyzed against buffer E^{NusG} (20 mM Tris/HCl (pH 7.5), 100 mM NH₄Cl, 10 mM MgCl₂, 0.5 mM EDTA, 1 mM tris(2-carboxyethyl)phosphine (TCEP)) overnight at 4 °C, before being subjected to a SEC run using a Superdex 200 column (GE Life Sciences). Fractions containing pure His₆-NusG were combined and stored at -80 °C.

NusG-NTD was produced as was full-length NusG for NMR studies, except that plasmid pET11a_*EcNusG-NTD(1-124)*, which encodes tag-less *E. coli* NusG-NTD (residues 1-124; (Burmam et al., 2011)), was used.

NusG-CTD was produced according to (Burmam, 2010) using *E. coli* BL21 (λ DE3) cells (Novagen, Madison, WI, USA) containing plasmid pETGB1a_*nusG-CTD(123-181)* (encoding *E. coli* NusG-CTD fused to a His₆-Gbl tag followed by a tobacco etch virus (TEV) cleavage site at its N-terminus). In brief, the conditions for expression were the same as for full-length NusG. Cells were collected by centrifugation (6,000 x g, 10 min, 4 °C), resuspended in buffer A^{NusG-CTD} (50 mM Tris/HCl (pH 7.5), 150 mM NaCl) containing 10 mM imidazole, supplemented with ½ protease inhibitor cocktail tablet (cOmplete, EDTA-free, Roche Diagnostics, Mannheim, Germany) and a small amount of DNase I (AppliChem GmbH, Darmstadt, Germany), and lysed with a microfluidizer. Upon centrifugation, the filtered (0.45 μ m) crude extract was loaded onto a 5 ml Ni²⁺-HiTrap column (GE Healthcare, Munich, Germany) buffer A^{NusG-CTD}. The column was washed with 20 CVs buffer A^{NusG-CTD} containing 10 mM imidazole and elution was carried out with a step gradient from 60 mM – 1 M imidazole in buffer A^{NusG-CTD}. Fractions containing the target protein were combined and dialyzed against buffer A^{NusG-CTD} (MWCO 3.5 kDa) at 4 °C overnight in the presence of TEV protease to cleave off the tag. The dialysate was applied to three coupled 5ml Ni²⁺-HiTrap columns (GE Healthcare, Munich, Germany) equilibrated with buffer A^{NusG-CTD} to remove His₆-Gbl, uncut fusion protein, and TEV protease. The flow through was concentrated by

ultrafiltration (MWCO 3 kDa) and then subjected to a size exclusion chromatography using a HiLoad S75 column (GE Healthcare, Munich, Germany) equilibrated with buffer B^{NusG-CTD} (25 mM HEPES (pH 7.5), 100 mM NaCl). Fractions containing pure NusG-CTD were combined and concentrated by ultrafiltration (MWCO 3 kDa), before being shock-frozen in liquid nitrogen and stored at -80 °C.

The production of the S10^Δ:NusB heterodimer was done according to (Zuber et al., 2019). For expression, *E. coli* BL21 (λ DE3) cells (Novagen, Madison, WI, USA) containing plasmids pGEX-6P_*ecoNusE*^Δ (encoding *E. coli* S10 where residues 46 - 67 were substituted by a single Ser (S10^Δ) with an N-terminal glutathione S-transferase (GST)-tag followed by a PreScission protease cleavage site) or pET29b_*ecoNusB* (encoding tag-less *E. coli* NusB) were grown in ampicillin (100 µg/ml) or kanamycin (30 µg/ml) containing LB medium, respectively, at 37 °C to an *OD*₆₀₀ of 0.5. The temperature was then lowered to 20 °C and over-expression was induced by addition of 0.5 mM IPTG after 30 min. Upon incubation overnight, cells were pelleted by centrifugation (6,000 x *g*). Cells containing S10^Δ or NusB, obtained from the same culture volume, were then resuspended in buffer A^{S10^Δ:NusB} (50 mM Tris/HCl (pH 7.5), 150 mM NaCl, 1 mM dithiothreitol (DTT)), supplemented with ½ protease inhibitor cocktail tablet (cOmplete, EDTA-free, Roche Diagnostics, Mannheim, Germany) and a small amount of DNase I (AppliChem GmbH, Darmstadt, Germany), mixed and subsequently lysed using a microfluidizer. The lysate was stirred for 30 min at 4 °C to allow formation of the S10^Δ:NusB dimer. Cell debris was removed by centrifugation and the crude extract was applied to four coupled 5 ml GSTrap FF columns (GE Healthcare, Munich, Germany) equilibrated with buffer A^{S10^Δ:NusB}. After washing with 20 column volumes (CVs) of buffer A^{S10^Δ:NusB} elution was performed in one step with buffer A^{S10^Δ:NusB} containing 15 mM reduced glutathione. PreScission protease was added to the combined target fractions and the protein solution was dialyzed against buffer B^{S10^Δ:NusB} (50 mM Tris/HCl (pH 7.5), 1 mM DTT)

overnight. The dialysate was subsequently applied to two 5 ml HiTrap Q XL columns coupled to two HiTrap SP XL columns (all from GE Healthcare, Munich, Germany) equilibrated with buffer $B^{S10\Delta:NusB}$. The columns were washed with 20 CVs buffer $B^{S10\Delta:NusB}$ and, after disconnecting, the S10 Δ :NusB dimer was eluted from the two HiTrap SP XL columns with buffer $B^{S10\Delta:NusB}$ containing 1 M NaCl. The protein solution was dialyzed against buffer $C^{S10\Delta:NusB}$ (25 mM HEPES (pH 7.5), 100 mM NaCl), concentrated *via* ultrafiltration (MWCO 5 kDa), flash-frozen in liquid nitrogen, and stored at -80 °C.

RNAP was produced as described (Zuber et al., 2019). Expression was carried out in *E. coli* BL21 (λ DE3) cells (Novagen, Madison, WI, USA) harboring plasmid pVS10 (encoding *E. coli* core RNAP subunits α , β , β' , and ω , with β carrying a C-terminal His₆-tag; (Svetlov and Artsimovitch, 2015)). Cells were grown at 37 °C in LB medium supplemented with ampicillin (100 μ g/ml) to an OD_{600} of 0.7. The temperature was lowered to 16 °C and gene expression was induced by addition of 0.5 mM IPTG at $OD_{600} = 0.8$. Cells were harvested by centrifugation (6,000 x g, 10 min, 4 °C) after overnight incubation, resuspended in buffer A^{RNAP} (50 mM Tris/HCl (pH 6.9), 500 mM NaCl, 5 % (v/v) glycerol, 1 mM β -mercaptoethanol (β -ME)) containing 10 mM imidazole, DNase I (AppliChem GmbH, Darmstadt, Germany), and 1/2 protease inhibitor tablet (cOmplete, EDTA-free, Roche Diagnostics, Mannheim, Germany) and lysed with a microfluidizer. The lysate was centrifuged for 30 min at 13,000 rpm and 4 °C, and, subsequently, the supernatant was applied to a 40 ml Ni²⁺-Chelating Sepharose column (GE Healthcare, Munich, Germany) equilibrated with buffer A^{RNAP} containing 10 mM imidazole. Upon washing with 25 CVs of buffer A^{RNAP} containing 10 mM imidazole RNAP was eluted using a gradient from 90 mM – 1 M imidazole in buffer A^{RNAP} . Fractions containing RNAP were combined, dialyzed against buffer B^{RNAP} (50 mM Tris/HCl (pH 6.9), 5 % (v/v) glycerol, 0.5 mM ethylenediaminetetraacetic acid (EDTA), 1 mM β -ME) containing 100 mM NaCl, and applied

to two 5 ml Heparin FF columns (GE Healthcare, Munich, Germany) equilibrated with buffer B^{RNAP} containing 100 mM NaCl. After washing with buffer B^{RNAP} containing 100 mM NaCl, the enzyme was eluted with a constant gradient from 100 mM to 1 M NaCl in buffer B^{RNAP}. RNAP-containing fractions were pooled, dialyzed against buffer C^{RNAP} (50 mM Tris/HCl (pH 6.9), 150 mM NaCl, 5 % (v/v) glycerol, 0.5 mM EDTA, 1 mM β -ME), and subsequently concentrated by ultrafiltration (MWCO 10 kDa) before being subjected to a SEC run using a HiLoad S200 column (GE Healthcare, Munich, Germany) equilibrated with buffer C^{RNAP}. Fractions containing pure core RNAP were concentrated by ultrafiltration (MWCO 10 kDa), glycerol was added to a final concentration of 50 % (v/v), and the solution was stored at -20 °C.

Intact 70S ribosomes were produced as follows. *E. coli* strain MRE600 cells grown in LB medium were harvested, lysed by passing through a French Press 3x at ~800 PSI, and clarified by a short centrifugation (20,000 rpm, 40 min) in opening buffer (20 mM Tris/HCl (pH 7.5), 100mM NH₄Cl, 10.5 mM Mg acetate, 0.5 mM EDTA, with half a protease inhibitor cocktail tablet (Roche, EDTA-free), and 1mM TCEP added just before use). The lysate was loaded onto the top of 5 mL sucrose cushion (20 mM Tris-HCl (pH 7.5), 500 mM NH₄Cl, 10.5 mM Mg acetate, 0.5 mM EDTA, 1.1 M sucrose, and 1 mM TCEP added before use) and centrifuged for 24 h at 28,000 rpm in a 70Ti rotor (Beckman Coulter, Inc.). The pellets were suspended in washing buffer (20 mM Tris-HCl (pH7.5), 500mM NH₄Cl, 10.5mM Mg acetate, 0.5 mM EDTA and 1 mM TCEP added before use), and centrifuged through a 10–35% (w/v) sucrose gradient for 19 h at 16,000 rpm in a SW28 rotor (Beckman Coulter, Inc.). Fractions containing the 70S ribosomes were pooled and kept at -80°C for further use.

Ribosomes for NMR experiments were obtained from New England Biolabs.

Electron Microscopy. Purified 70S ribosomes were incubated with full-length NusG at a ratio of 1:7 for 40 min at room temperature, prior to blotting and plunge-freezing as

previously described (Grassucci et al., 2007). Data were collected on a TF30 Polara electron microscope (FEI, Portland, Oregon) at 300kV using a K2 Summit direct electron detector camera (Gatan, Pleasanton, CA). Images were recorded using the automated data collection system Leginon (Suloway et al., 2005) in counting mode, and taken at the nominal magnification of 32,000x, corresponding to a calibrated pixel size of 1.66Å.

Image processing. A total of 188,127 particles were automatically extracted from 1,327 images using Arachnid (Langlois et al., 2014). RELION (Scheres, 2012) 3D classification was used to resolve the heterogeneity of the particle images, and auto-refinement to further improve resolution for each class. The final refinement for the NusG-bound 70S class containing 17,122 particles yielded an average resolution of ~6.8Å (FSC=0.143; following “gold standard” protocol, see table S1).

Model refinement. The model refinement was performed in two stages; the first was molecular dynamics flexible fitting and the second was fine-tuning of the model using the real-space refinement function in Phenix.

In the first stage, the starting model was assembled from the X-ray structure of the *E. coli* 30S ribosomal subunit (PDB ID 4GD2) and the NMR solution structure of the NusG-CTD (PDB 2KVQ chain G). This starting model was first docked into the segmented maps of our 70S density map as a rigid body using UCSF Chimera (Pettersen et al., 2004). Then it was fitted into the segmented map using the Molecular Dynamic Flexibly Fitting (MDFF) method (Trabuco et al., 2008) and run using the NAMD program (Phillips et al., 2005) for 0.5 ns of simulation time, followed by 5,000 steps of energy minimization.

In the second stage, we performed rounds of real-space refinement using the program Phenix (Afonine et al., 2018) to correct geometry, rotamers and overlaps. Atomic positions from the model obtained by MDFF were also used as the reference model for restraints used

during the refinement to retain the secondary structure. To account for the relatively low resolution of the map, the parameters of *weight* and *nonbonded_weight* for the restraint terms were manually adjusted. The model was inspected, and problematic outliers were fixed using the program Coot (Emsley et al., 2010). The final model was validated using the program MolProbity (Williams et al., 2018).

NMR spectroscopy. NMR experiments were conducted on Bruker Ascend Aeon 900 and 1000 MHz spectrometers equipped with cryogenically cooled, inverse triple resonance probes at 298 K. NMR data was converted and processed using in-house software. 2D correlation spectra were visualized and analyzed with NMRViewJ (One Moon Scientific, Inc., Westfield, NJ, USA), 1D spectra were plotted using MatLab (The MathWorks, Inc., Version 9.2.0.538062). Resonance assignments for NusG methyl groups were taken from a previous study (Mooney et al., 2009).

[ILV]-NusG-CTD was in 10 mM K-phosphate (pH 7.5), 50 mM KCl, 1 mM MgCl₂, 99.9 % (v/v) D₂O, [ILV]-NusG-NTD in 50 mM Na-phosphate (pH 7.5), 50 mM KCl, 0.3 mM EDTA, 5 % (v/v) d₇-glycerol, 0.01 % (w/v) NaN₃, 99.9 % (v/v) D₂O. For the titration of [ILV]-NusG with RNAP and S10^Δ-NusB, all proteins were in 50 mM Na-phosphate (pH 7.5), 50 mM KCl, 0.3 mM EDTA, 99.9 % (v/v) D₂O and 5 mM MgCl₂ and 2 mM DTT were added to the NMR sample to increase the long-term stability of RNAP. For all interaction studies involving ribosomes and for the titration of [ILV]-NusG^{F165A} with S10^Δ-NusB, all components were in 20 mM HEPES/KOH (pH 7.6), 10 mM Mg-acetate, 30 mM KCl, 7 mM β-ME, 10 % (v/v) D₂O. The titration of [ILV]-NusG^{F165A} with S10^Δ-NusB was conducted in a 5 mm tube with an initial sample volume of 550 μl. All other measurements were carried out in 3 mm NMR tubes with an (initial) volume of 200 μl.

1D and 2D titration experiments were evaluated quantitatively by analyzing either changes in signal intensity or changes in chemical shifts. If chemical shift changes were in the fast regime on the chemical shift the normalized chemical shift perturbation ($\Delta\delta_{\text{norm}}$) was calculated according to equation 1.

$$\Delta\delta_{\text{norm}} = \sqrt{(\Delta\delta^{\text{1H}})^2 + [0.25 (\Delta\delta^{\text{13C}})]^2} \quad (1)$$

with $\Delta\delta$ being the resonance frequency difference between the initial and final state of the titration (i.e. [ILV]-NusG:RNAP:S10:NusB = 1:2:0:0 vs. 1:2:2:2) in ppm.

If the system was in slow or intermediate chemical exchange the signal intensities were analyzed quantitatively. First, the intensity of each 1D spectrum or methyl group signal, respectively, was normalized by the concentration of the [ILV]-labeled protein, the receiver gain, the number of scans, and the length of the 90° ¹H pulse. Then the relative intensity, i.e. the ratio of the normalized signal intensity of [ILV]-labeled protein in the respective titration step to the normalized signal intensity of free [ILV]-labeled protein, was calculated and plotted against the sequence of NusG or the NusG variant, respectively.

qRT-PCR. Total RNA was extracted from cells grown in M9 medium supplemented with casamino acids (0.2%) at 37°C to mid-log phase ($OD_{600}=0.3$). Fold-increase of the PCR product was determined using qRT-PCR. RNA was extracted from logarithmically growing cultures ($OD_{600}=0.2-0.3$). Where indicated, cells were treated with BCM (100 mg/ml) 1 min before induction with 1mM IPTG for *lacZ*. Samples were removed (0.5ml) at the indicated times and total RNA extracted using Qiagen RNeasy and Qiagen RNa protect Bacteria Reagent (Qiagen, Germantown, MD). cDNA was synthesized from the samples using High Capacity RNA to cDNA kit (ThermoFisher, Waltham, MA). qRT-PCR reactions

were performed using Taqman Gene Expression Master Mix (Thermofisher, Waltham, MA) and Biorad DNA Engine Opticon2 Real-Time Cycler (Bio-Rad Laboratories, Hercules, CA) and PrimeTime qPCR probes (Integrated DNA Technologies, Coralville, IA).

The *lacA* transcript was probed with the following probe:

5'-/56-FAM/CCACATGAC/ZEN/TTCCGATCCAGACGTT/3IABkFQ/-3';

primer 1: 5'-ATACTACCCGCGCCAATAAC;

primer 2: 5'-CCCTGTACACCATGAATTGAGA).

The reference gene was *ompA*

(probe: 5'-/56-FAM/CAACAACAT/ZEN/CGGTGACGCACACAC /3IABkFQ/-3';

primer 1: 5'-TGACCGAAACGGTAGGAAAC;

primer 2: 5'-ACGCGATCACTCCTGAAATC).

The PCR was performed using the following conditions: 50 °C for 10 min., 95 °C for 2 min, followed by 40 cycles each of 95 °C for 15s, and 60 °C 1min; 50 nmol probe, 25 nmol primer 1, 25 nmol primer 2. Fold increases were calculated from measured C_t values using the $\Delta\Delta C_t$ method (Livak and Schmittgen, 2001). Read-through was calculated from the ratio of fold-increase of RNA level +/- BCM. Values are the average of three or more independent experiments and all reactions were performed in duplicate. The standard deviation of read-through was calculated using error propagation.

β -galactosidase assays. Cultures were grown in LB to early log phase ($OD_{600} = 0.3$) at 37 °C. Where indicated BCM (100 mg/ml) was added to inhibit Rho-dependent transcription termination prior to induction of *lacZ* with 1mM IPTG. Where indicated λn was expressed by incubation at 42 °C. Reactions were terminated 15 min after induction. β -galactosidase was measured using a modification of the method of Miller (Zhang and Bremer, 1995). Read-

through was calculated from the ratio of β -galactosidase activity \pm BCM/ λ N. At least three replicates were performed per experiment and the resultant values were averaged. The standard deviation of read-through was calculated using error propagation.

Supplemental References

- Afonine, P.V., Poon, B.K., Read, R.J., Sobolev, O.V., Terwilliger, T.C., Urzhumtsev, A., and Adams, P.D. (2018). Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr D Struct Biol* 74, 531-544.
- Burmann, B.M., Scheckenhofer, U., Schweimer, K., and Rösch, P. (2011). Domain interactions of the transcription-translation coupling factor *Escherichia coli* NusG are intermolecular and transient. *Biochem. J.* 435, 783–789.
- Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66, 486-501.
- Grassucci, R.A., Taylor, D.J., and Frank, J. (2007). Preparation of macromolecular complexes for cryo-electron microscopy. *Nat Protoc* 2, 3239–3246.
- Langlois, R., Pallesen, J., Ash, J.T., Nam Ho, D., Rubinstein, J.L., and Frank, J. (2014). Automated particle picking for low-contrast macromolecules in cryo-electron microscopy. *J. Struct. Biol.* 186, 1–7.
- Livak, K.J., and Schmittgen, T.D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* 25, 402–408.
- Meyer, O., and Schlegel, H.G. (1983). Biology of aerobic carbon monoxide-oxidizing bacteria. *Annu. Rev. Microbiol.* 37, 277–310.
- Mooney, R.A., Schweimer, K., Rösch, P., Gottesman, M., and Landick, R. (2009). Two structurally independent domains of *E. coli* NusG create regulatory plasticity *via* distinct interactions with RNA polymerase and regulators. *J. Mol. Biol.* 391, 341–358.
- Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera--a visualization system for exploratory research and analysis. *J. Comput. Chem.* 25, 1605–1612.
- Phillips, J.C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel,

R.D., Kalé, L., and Schulten, K. (2005). Scalable molecular dynamics with NAMD. *J. Comput. Chem.* *26*, 1781–1802.

Sambrook, J., and Russel, D.W. (2001). *Molecular Cloning: A Laboratory Manual* (New York: Cold Spring Harbor Press).

Saxena, S., Myka, K.K., Washburn, R., Costantino, N., Court, D.L., and Gottesman, M.E. (2018). *Escherichia coli* transcription factor NusG binds to 70S ribosomes. *Mol. Microbiol.* *108*, 495–504.

Scheres, S.H.W. (2012). RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* *180*, 519–530.

Sharan, S.K., Thomason, L.C., Kuznetsov, S.G., and Court, D.L. (2009). Recombineering: a homologous recombination-based method of genetic engineering. *Nat. Protoc.* *4*, 206–223.

Silhavy, T.J., Berman, M.L., and Enquist, L.W. (1984). *Experiments with Gene Fusions* (Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press).

Suloway, C., Pulokas, J., Fellmann, D., Cheng, A., Guerra, F., Quispe, J., Stagg, S., Potter, C.S., and Carragher, B. (2005). Automated molecular microscopy: the new Legimon system. *J. Struct. Biol.* *151*, 41–60.

Svetlov, V., and Artsimovitch, I. (2015). Purification of bacterial RNA polymerase: tools and protocols. *Methods Mol. Biol.* *1276*, 13–29.

Trabuco, L.G., Villa, E., Mitra, K., Frank, J., and Schulten, K. (2008). Flexible fitting of atomic structures into electron microscopy maps using molecular dynamics. *Structure* *16*, 673–683.

Williams, C.J., Headd, J.J., Moriarty, N.W., Prisant, M.G., Videau, L.L., Deis, L.N., Verma, V., Keedy, D.A., Hintze, B.J., Chen, V.B., *et al.* (2018). MolProbity: More and better reference data for improved all-atom structure validation. *Protein Sci* *27*, 293–315.

Zhang, X., and Bremer, H. (1995). Control of the *Escherichia coli* *rrnB* P1 promoter strength by ppGpp. *J. Biol. Chem.* *270*, 11181–11189.

Zuber, P.K., Schweimer, K., Rösch, P., Artsimovitch, I., and Knauer, S.H. (2019). Reversible fold-switching controls the functional cycle of the antitermination factor RfaH. *Nat Commun* 10, 702.