

NusG inhibits RNA polymerase backtracking by stabilizing the minimal transcription bubble

Matti Turtola, Georgiy A Belogurov*

Department of Biochemistry, University of Turku, Turku, Finland

Abstract Universally conserved factors from NusG family bind at the upstream fork junction of transcription elongation complexes and modulate RNA synthesis in response to translation, processing, and folding of the nascent RNA. *Escherichia coli* NusG enhances transcription elongation in vitro by a poorly understood mechanism. Here we report that *E. coli* NusG slows Gre factor-stimulated cleavage of the nascent RNA, but does not measurably change the rates of single nucleotide addition and translocation by a non-paused RNA polymerase. We demonstrate that NusG slows RNA cleavage by inhibiting backtracking. This activity is abolished by mismatches in the upstream DNA and is independent of the gate and rudder loops, but is partially dependent on the lid loop. Our comprehensive mapping of the upstream fork junction by base analogue fluorescence and nucleic acids crosslinking suggests that NusG inhibits backtracking by stabilizing the minimal transcription bubble.

DOI: [10.7554/eLife.18096.001](https://doi.org/10.7554/eLife.18096.001)

Introduction

To control when and how fast genes are expressed, the activity of RNA polymerase (RNAP) is tightly regulated. Much of transcription regulation in all domains of life takes place at the initiation stage by modulating activities of promoters. The examples of on/off regulation at the transcript elongation stage, such as promoter-proximal pauses in metazoans (Adelman and Lis, 2012) and antitermination in prokaryotes (Santangelo and Artsimovitch, 2011), are also known. In other cases, transcription elongation control is mediated by coupling of transcription to downstream processes, such as RNA translation, processing, and folding (Proshkin et al., 2010; Bubunenko et al., 2013). The multisubunit RNAPs evolved to elongate relatively inefficiently in the absence of proper coupling, thereby enabling the downstream processes to control the elongation rate. The ubiquitous family of NusG proteins (SPT5/SPT4 in archaea and yeast, DSIF in mammals) are the central components which mediate coupling between transcription and the downstream processes (Werner, 2012).

The simplest member of the family, a bacterial NusG, consists of two domains connected by a flexible linker (Figure 1). The N-terminal domain (NTD or NGN) anchors NusG to the clamp helices of the RNAP β' subunit, whereas the C-terminal domain (CTD or KOW) interacts with the components of the downstream processes (reviewed in [Belogurov and Artsimovitch, 2015]). In *E. coli*, NusG CTD interacts with NusE as part of the ribosome on protein coding operons (Burmam et al., 2010) or as a part of a so-called antitermination complex (NusABEG) on ribosomal RNA operons (Zellers and Squires, 1999; Shankar et al., 2007; Singh et al., 2016); in these contexts, NusG inhibits the function of a transcription termination factor Rho. If neither ribosome nor antitermination complex is engaged, which often implies that transcription is futile, NusG CTD binds to Rho and facilitates termination of transcription (Cardinale et al., 2008; Peters et al., 2012). tRNA and other non-coding RNA genes escape the premature termination by Rho possibly due to their extensive secondary structures and small size relative to the transcribed region required for the termination of

*For correspondence: gebelo@utu.fi

Competing interests: The authors declare that no competing interests exist.

Funding: See page 23

Received: 23 May 2016

Accepted: 03 October 2016

Published: 04 October 2016

Reviewing editor: Robert Landick, University of Wisconsin-madison, United States

© Copyright Turtola and Belogurov. This article is distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use and redistribution provided that the original author and source are credited.

eLife digest Cells decode genes in two steps. First, they synthesize a molecule similar to DNA, called RNA, which is a complementary copy of the gene. This process, known as transcription, creates an intermediate RNA molecule that is turned into protein in the second step. RNA polymerase is an enzyme that carries out transcription; it separates the two strands of the DNA helix so that the RNA can be synthesized from the DNA template. By opening up the DNA downstream of where active copying is taking place, and re-annealing it upstream, RNA polymerase maintains a structure called a "transcription bubble". RNA polymerases do not copy continuously but oscillate back and forth along the DNA. Sometimes larger backwards oscillations, known as backtracking, temporarily block the production of the RNA molecule and slow down the transcription process.

A protein called NusG helps to couple transcription to the other related processes that happen at the same time. One end of the protein, the N-terminal domain, anchors it to RNA polymerase and stimulates transcription elongation. The other end, the C-terminal domain, interacts with other proteins involved in the related processes and can positively or negatively control transcription elongation. Nevertheless it was poorly understood how NusG carries out these roles.

Turtola and Belogurov investigated how NusG from the bacterium *Escherichia coli* affects the individual steps of transcription elongation. A simple experimental system was used, consisting of short pieces of DNA and RNA, an RNA polymerase and NusG. A transcription bubble resembles an opening in a zipper with two sliders; and rather than affecting the synthesis of RNA, NusG affected the part that corresponds to the "slider" located at the rear edge of the bubble. NusG helped this slider-like element to bring the DNA strands at this edge of the bubble back together and modified it so that it behaved as a ratchet that inhibited RNA polymerase from backtracking. This did not affect the smaller backwards and forwards oscillations of RNA polymerase.

Turtola and Belogurov suggest that these newly discovered effects play a key role in regulating transcription; NusG's N-terminal domain makes the RNA polymerase more efficient, whilst the C-terminal domain makes it amenable to control by other proteins. Future studies will investigate whether these effects are seen in more complex experimental systems, which include proteins that interact with NusG.

DOI: [10.7554/eLife.18096.002](https://doi.org/10.7554/eLife.18096.002)

transcription by Rho (**Mooney et al., 2009a; Peters et al., 2009**). NusG-mediated coupling of transcription with the pioneer round of translation is likely conserved in prokaryotes, whereas functioning of NusG CTD (and additional KOW domains present in eukaryotic SPT5 and DSIF) in RNA processing/maturation is likely conserved in all domains of life (**Belogurov and Artsimovitch, 2015**).

The regulation of gene expression by NusG-like proteins, which include several paralogs in some bacterial species, is complex. Even the housekeeping NusG may exhibit opposite effects on transcription in vivo depending on the protein partner(s) bound to its CTD domain. Furthermore, NusG from different bacteria display seemingly opposite effects on transcription by their cognate RNAPs in a purified in vitro system lacking the downstream components. The *E. coli* NusG has an intrinsic stimulatory effect on transcript elongation in vitro (**Bar-Nahum et al., 2005; Burova et al., 1995**), which persists when an isolated NusG NTD is used (**Mooney et al., 2009b**). It is hypothesized that this intrinsic stimulatory effect of NusG NTD may allow RNAP to transcribe more efficiently in vivo when coupled with the downstream processes and slower if the coupling is broken (**McGary and Nudler, 2013; Belogurov and Artsimovitch, 2015; Burmann and Rösch, 2011**). However, NusG from *Thermus thermophilus* slows down its cognate RNAP in vitro (**Sevostyanova and Artsimovitch, 2010**), whereas *Bacillus subtilis* NusG stimulates pausing by interacting with specific sequences in the non-template DNA (**Yakhnin et al., 2008, 2016**). We later suggest that these apparent incongruences result from the superimposition of several distinct consequences of the NusG NTD binding to the RNAP and considerable differences in the elongation properties of these RNAPs. But first we consider the mechanistic details of the elongation stimulation by *E. coli* NusG.

The RNAP nucleotide addition cycle consists of (i) the NTP substrate binding to a post-translocated transcription elongation complex (TEC); (ii) a chemical step of the nucleotide incorporation; and (iii) the post-catalytic relaxation of the resulting pre-translocated TEC, which involves the release

of pyrophosphate and forward translocation (reviewed in [Belogurov and Artsimovitch, 2015]). The processive repetition of this cycle is sometimes interrupted by off-pathway pause events. The latter can be classified into pauses involving backtracking of the RNAP on the DNA template by two or more registers, which occludes the active site with the nascent RNA (Komissarova and Kashlev, 1997; Nudler et al., 1997), and diverse non-backtracked pauses (Artsimovitch and Landick, 2000). The non-backtracked pauses involve more complex and less understood rearrangement of the active site and the RNAP structure that impede the nucleotide addition (Hein et al., 2014; Zhang et al., 2010; Kireeva and Kashlev, 2009). Notably, many non-backtracked pauses likely involve partial opening of the β' clamp (Hein et al., 2014; Weixlbaumer et al., 2013), a large mobile domain that contributes most of the β' subunit contacts with the nucleic acids.

The elongation-enhancing effect of NusG may arise from (i) accelerating the on-pathway elongation, (ii) diminishing some or all type of pauses, or both. Early studies suggested that NusG acts by reducing pausing (Burova et al., 1995). Recent reports further specified that NusG homologues enhance elongation by restricting the conformational flexibility of the RNAP β' clamp (Sevostyanova et al., 2011; Hirtreiter et al., 2010), which is consistent with biophysical measurements (Schulz et al., 2016) and structural considerations (Martinez-Rucobo et al., 2011; Klein et al., 2011). On the other hand, *E. coli* NusG mainly reduces the frequency and duration of backtracked pauses (Herbert et al., 2010) that are not explicitly known to involve the β' clamp

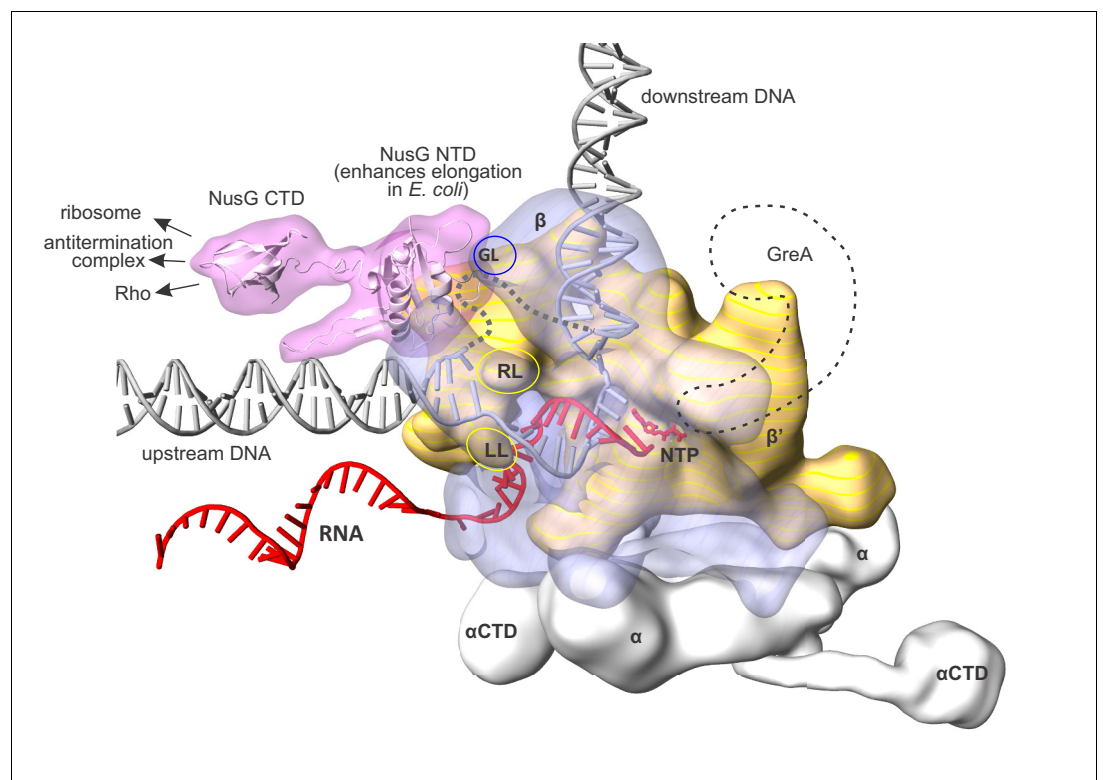


Figure 1. An overview of the bacterial transcription elongation complex (TEC) with bound NusG. RNAP core subunits are depicted by simplified differentially colored molecular surfaces. β is depicted transparent to reveal the path of nucleic acids through the enzyme. The positions of NusG CTD and α CTD, connected via flexible linkers, were chosen arbitrary. The locations of RNAP cleft loops individually deleted in this study, β Gate Loop (GL), β' Rudder Loop (RL) and β' Lid Loop (LL) are accentuated by ovals. The hypothetical path of the single stranded non-template DNA is depicted by a grey dashed line. The approximate location of GreA cleavage factor employed in backtracking experiments (see results) is depicted as a black dashed contour. The composite model was generated using the *Thermus thermophilus* TEC (Vassilyev et al., 2007), NusG NTD from the NusG-RNAP model in Martinez-Rucobo et al., 2011 and the elements from other structures (see Materials and methods). The duplex DNA immediately upstream of the RNA:DNA hybrid was modeled *de novo* as described in the Results section.

DOI: 10.7554/eLife.18096.003

opening (Sekine *et al.*, 2015) and has only small stimulatory effect at non-backtracked pauses (Artsimovitch and Landick, 2000; Belogurov *et al.*, 2010; Kolb *et al.*, 2014).

Here, we report a detailed analysis of NusG effects on the individual steps of the nucleotide addition cycle, backtracking, and the conformation of the upstream DNA. Our results suggest that NusG slows backtracking without affecting the on-pathway elongation in non-paused TECs. We also demonstrate that NusG inhibits backtracking by restricting the melting of the upstream DNA, independently of the NusG-RNAP contacts that are important for stabilization of the β' clamp conformation. We further perform a comprehensive mapping of the upstream fork junction, determine the point of the upstream DNA reannealing, and provide a plausible mechanistic model of the anti-backtracking action of NusG.

Results

NusG does not affect the rates of nucleotide incorporation and translocation in a non-paused TEC

To determine the effect of NusG on the kinetics of nucleotide addition and translocation, we utilized a TEC design that was extensively validated in our previous studies (Malinen *et al.*, 2012, 2015). The TEC was assembled on a synthetic nucleic acid scaffold and contained the fully complementary transcription bubble flanked by 20-nucleotide DNA duplexes upstream and downstream (Figure 2—figure supplement 1). The annealing region of a 16-nucleotide RNA primer was initially nine nucleotides, permitting the TEC extended by one nucleotide to adopt the post- and pre-translocated states, but disfavoring backtracking. The RNA primer was 5' labeled with the infrared fluorophore ATTO680 to monitor the RNA extension by the denaturing PAGE. We performed parallel, time-resolved measurements of nucleotide incorporation by assembled TEC accompanied by forward translocation and, in a separate experiment, pyrophosphorolysis of extended TEC accompanied by backward translocation in the presence and absence of saturating concentration of NusG (2 μ M, see later) (Figure 2A–C). RNA extension was monitored by a rapid chemical quench-flow method, whereas forward and backward translocation were monitored by measuring the fluorescence of the 6-methyl-isoxanthopterin base (6-MI) incorporated in the template DNA strand in a stopped-flow instrument (Malinen *et al.*, 2015). NusG did not affect either of these on-pathway reactions. We then tested the effect of NusG on the TEC translocation bias in the equilibrium setup (Figure 2D). We have previously demonstrated that the predominantly post-translocated TEC can be converted into the pre-translocated state by tagetitoxin (TGT) (Malinen *et al.*, 2012). TGT was equally potent in biasing the TEC towards the pre-translocated state in the presence and absence of NusG, suggesting that NusG does not affect the equilibrium between the post- and pre-translocated states. Overall, these experiments suggest that NusG does not measurably affect the on-pathway kinetics and thermodynamics of transcript elongation (Table 1) in the non-paused TECs examined in this study.

NusG slows backtracking captured by GreA-mediated RNA cleavage

NusG has been suggested to inhibit stochastic and sequence-specific backtracking during transcription through a long template in vitro (Herbert *et al.*, 2010). We developed a system where backtracking of the TEC is driven by the rapid cleavage of the nascent RNA by the RNAP active site. The reaction was initiated by adding 2–8 μ M GreA protein that transforms the RNAP active site into a highly efficient RNase. The TECs were assembled on a synthetic nucleic acid scaffold and contained the fully complementary transcription bubble (Figure 2—figure supplement 1). The annealing region of 18-nucleotide long RNA primer was 11 nucleotides, allowing the TEC to backtrack by one nucleotide. The RNA primer was 5' labeled with ATTO680 to monitor the accumulation of a 16-nucleotide RNA cleavage product in a rapid chemical quench flow experiment. The RNA primer also contained 2-aminopurine (2-AP) as the penultimate 3' nucleotide (Figure 3A), thereby permitting monitoring of 2-AP-p-C dinucleotide release by measuring an increase in 2-AP fluorescence in a stopped flow instrument. In a subset of experiments, the template DNA contained 6-MI nine registers upstream of the RNA 3' end (Figure 3A) to directly monitor RNAP backtracking by measuring the decrease in 6-MI fluorescence in a stopped flow instrument. Importantly, the changes in 6-MI and 2-AP fluorescence were driven by the RNA cleavage and not GreA binding because the addition

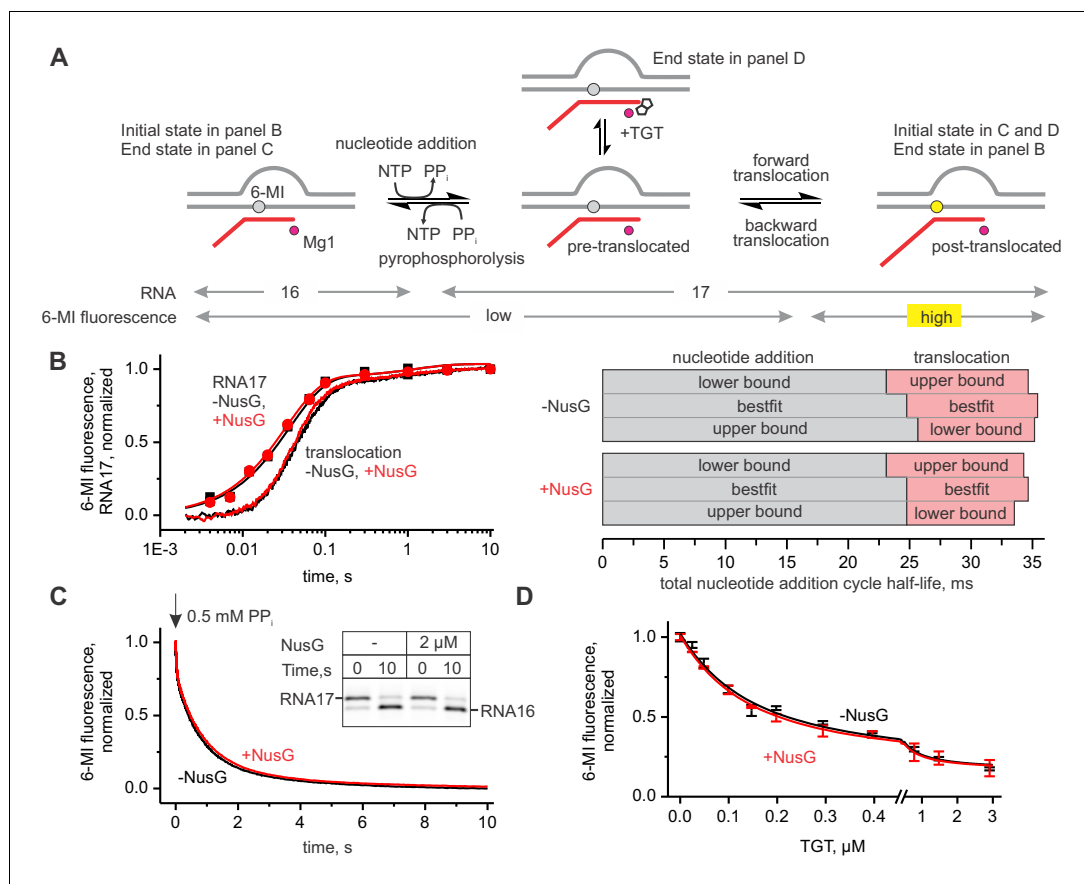


Figure 2. NusG does not affect on-pathway transcription elongation. (A) The schematics of the system used for monitoring the forward and the backward kinetics of the nucleotide addition cycle. In addition, the translocation bias of the TEC was evaluated under equilibrium conditions by measuring the TEC response to tagetitoxin (TGT). (B) The effect of NusG on the rate of nucleotide addition (discrete time-points and best-fit lines) and forward translocation (continuous time-traces). The lower and upper bounds of the reaction half-lives were calculated by combined analysis of data from several independent experiments (Table 5) by FitSpace routine of Kintek Explorer software (at a 10% increase in Chi²). (C) The effect of NusG on the pyrophosphorolysis-driven backward translocation. *Inset*: the gel control of the pyrophosphorolysis reaction. (D) The effect of NusG on the potency of TGT to bias the TEC towards the pre-translocated state. Error bars indicate the range of duplicate measurements. Numerical values of the reaction rate constants and the median reaction times are presented in Table 1.

DOI: 10.7554/eLife.18096.004

The following figure supplement is available for figure 2:

Figure supplement 1. TEC systems employed in the experiments presented in the main figures.

DOI: 10.7554/eLife.18096.005

of the cleavage-deficient D41N GreA variant did not change either fluorescence (Figure 3—figure supplement 1).

Parallel experiments demonstrated that within the limits of experimental uncertainty accumulation of the 16-nucleotide cleavage product, the release of 2AP-p-C dinucleotide, and RNAP backtracking occur with the same rate (Figure 3B). Importantly, 3' mismatched TECs that are more prone to backtracking were cleaved approximately thirty times faster than the matched TECs (Figure 3C), suggesting that for the matched TECs the rates of RNA cleavage and the dinucleotide release were limited by the rate of backtracking. We therefore concluded that all three types of measurements can be used interchangeably to monitor backtracking in this system.

Saturating concentrations of NusG (2 μM, see below) slowed RNA cleavage, dinucleotide release (increase in 2AP fluorescence), and backtracking (decrease in 6-MI fluorescence) approximately two-fold (Figure 3B). Similar inhibition was observed at 2 μM (Figure 3—figure supplement 3) and 8

Table 1. Numerical values of the reaction rate constants and the median reaction times.**Figure**

2B	–NusG			+NusG		
	lower bound	best fit	upper bound	lower bound	best fit	upper bound
nucleotide addition, s ⁻¹	27	28	30	28	28	30
translocation, s ⁻¹	60	65	73	62	70	79
Slow TEC fraction		~8%			~7%	
recovery rate, s ⁻¹	0.4	1.1	2.7	0.4	1.1	2.7
inactivation rate, s ⁻¹	0.03	0.09	0.3	0.02	0.08	0.2
The lower and upper bounds of rate constants were calculated by the combined analysis of data from several independent experiments (Table 5) by FitSpace routine of Kintek Explorer software (at a 10% increase in Chi2).						
2C	median pyrophosphorolysis time, s					
	-NusG			+NusG		
	0.49 ± 0.08			0.51 ± 0.08		
Errors indicate the range of the bestfit estimates in duplicate experiments.						
2D	K _D TGT, μM					
	-NusG			+NusG		
	0.09–0.15			0.09–0.14		
The ranges represent 95% confidence interval for K _D determined by the nonlinear regression analysis of data from two independent experiments.						
3B	Method					
	–NusG			+NusG		
	6-MI			28.0 ± 0.7		
	RNA18			30.4 ± 3.4		
	RNA16			30.1 ± 3.6		
	2-AP			30.3 ± 2.2		
Errors indicate the range of the bestfit estimates in duplicate experiments.						
3C				median reaction time, s		
				–NusG		
	RNA	DNA	RNAP	+NusG		
	matched	matched	WT	12.4 ± 1.8		
	matched	matched	ΔRL	11.1 ± 1.9		
	matched	matched	ΔLL	11.4 ± 1.3		
	matched	matched	ΔGL	24.0 ± 2.3		
	matched	mm 1	WT	19.9 ± 2.2		
	matched	mm 1 and 2	WT	2.30 ± 0.04		
	3' mm	matched	WT	0.34 ± 0.17		
3' mm stands for mismatch against 3' RNA nucleotide. Errors indicate the range of the bestfit estimates in duplicate experiments.						

3' mm stands for mismatch against 3' RNA nucleotide. Errors indicate the range of the bestfit estimates in duplicate experiments.
DOI: 10.7554/eLife.18096.006

μM GreA (**Figure 3B**), suggesting that NusG did not act by weakening the binding of GreA. Furthermore, NusG did not slow the dinucleotide release from the TEC that was biased towards the backtracked state by the RNA 3' end mismatch (**Figure 3C**), suggesting that NusG did not directly inhibit RNA cleavage.

NusG anti-backtracking activity requires the double-stranded upstream DNA

NusG may slow backtracking of the TEC by affecting the conformation of RNAP and/or transcription bubble. To dissect the mechanism of anti-backtracking activity of NusG, we individually deleted β

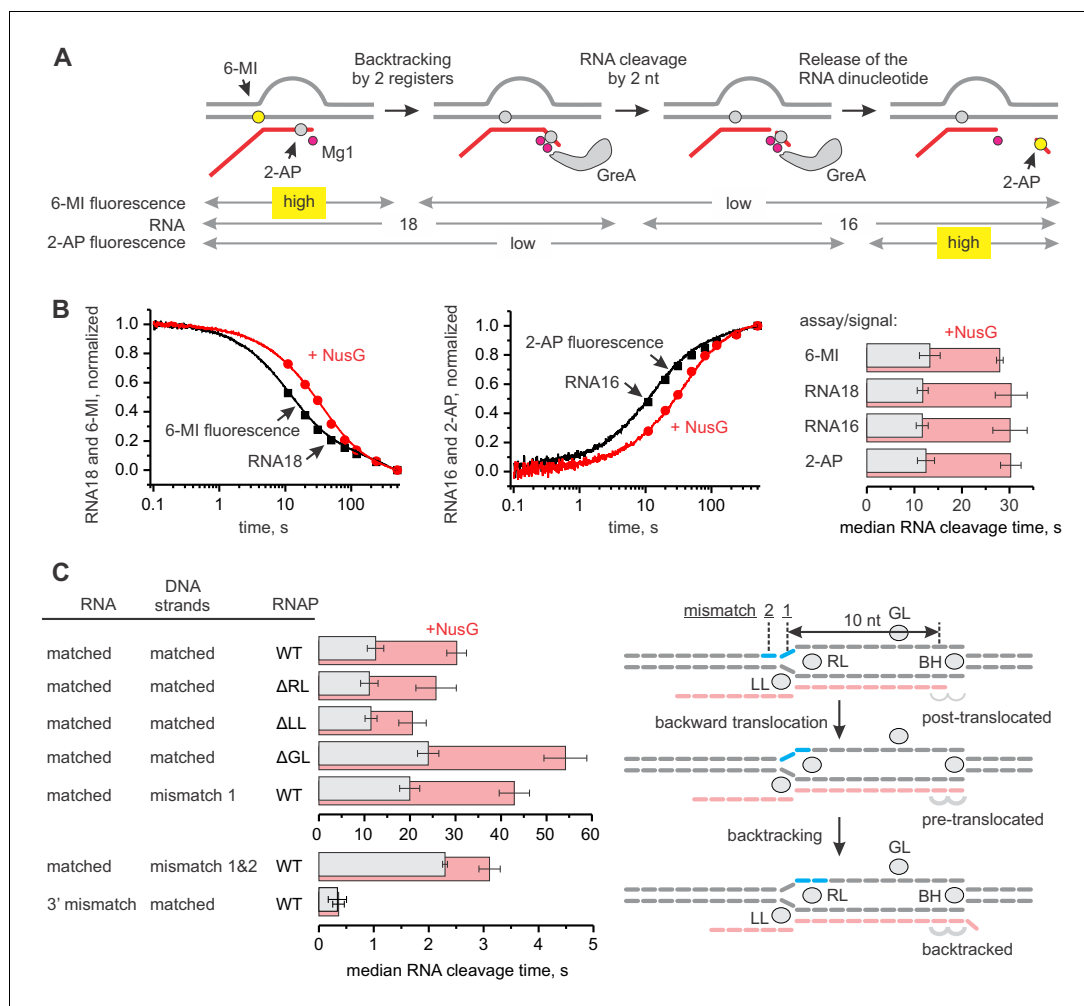


Figure 3. NusG inhibits GreA assisted RNA cleavage by slowing backtracking. (A) Three assays for monitoring the GreA-assisted RNA cleavage: TEC backtracking (6-MI fluorescence decrease), RNA cleavage (PAGE), and the dinucleotide release (2-AP fluorescence increase). (B) Left: the effect of NusG on the TEC backtracking (continuous time-traces) and RNA cleavage (discrete time-points) upon the addition of 8 μ M GreA. Center: the effect of NusG on the release of the cleaved dinucleotide (continuous time-traces) and the RNA cleavage (discrete time-points). Right: the median reaction times. (C) The effect of NusG on the RNA cleavage by TECs with deletions of the RNAP domains, the mismatched upstream DNA and a 3' rCda RNA:DNA mismatch. The median reaction times were determined by monitoring the increase in 2AP fluorescence at 8 μ M GreA. The schematic on the right illustrates the location of DNA:DNA mismatches (cyan bars). Numerical values of the median reaction times are presented in **Table 1**. Error bars indicate the range of the bestfit estimates in duplicate experiments.

DOI: [10.7554/eLife.18096.007](https://doi.org/10.7554/eLife.18096.007)

The following figure supplements are available for figure 3:

Figure supplement 1. Control experiment with the inactive GreA (D41N variant).

DOI: [10.7554/eLife.18096.008](https://doi.org/10.7554/eLife.18096.008)

Figure supplement 2. Primary data for graphs in **Figure 3B**: The effect of NusG on GreA assisted RNA cleavage.

DOI: [10.7554/eLife.18096.009](https://doi.org/10.7554/eLife.18096.009)

Figure supplement 3. The effect of NusG on RNA cleavage by wild-type TECs at 2 μ M GreA.

DOI: [10.7554/eLife.18096.010](https://doi.org/10.7554/eLife.18096.010)

and β' cleft loops near the NusG binding site (**Figure 1**). We evaluated the NusG effects on backtracking of TECs assembled with RNAPs lacking β Gate Loop (Δ GL), β' Lid Loop (Δ LL) and β' Rudder Loop (Δ RL). We also perturbed the base pairs of the upstream DNA that may affect backtracking. While the register of the upstream DNA reannealing is not exactly known, the bacterial TEC

structures suggest that (i) DNA may reanneal as early as ten nucleotides upstream of the RNA 3' end in the post-translocated TEC and (ii) DNA must be unpaired up to at least 11 nucleotides upstream of the RNA 3' end in one-nucleotide backtracked TEC (**Figure 3C**). Accordingly, we evaluated the anti-backtracking activity of NusG on the TECs where DNA reannealing ten and ten-eleven nucleotides upstream from the RNA 3' end was inhibited by mismatches.

Δ RL and Δ LL did not affect backtracking rates in the absence of NusG, but the latter deletion reduced the TEC response to NusG twofold (**Figure 3C**). In contrast, Δ GL and the DNA:DNA mismatch ten nucleotides upstream of the RNA 3' end slowed backtracking 1.5–2 fold but did not affect the TEC responses to NusG. Most notably, TEC with two DNA:DNA mismatches 10–11 nucleotides upstream of the RNA 3' end backtracked approximately fivefold faster than the fully-matched TEC and was insensitive to NusG (**Figure 3C**). At the same time, TEC with a 3' RNA:DNA mismatch cleaved RNA further sevenfold faster than the TEC with two DNA:DNA mismatches (**Figure 3C**), suggesting that backtracking limited the cleavage rate in the latter TEC. These results suggested that: (i) NusG slows backtracking by inhibiting DNA melting eleven nucleotides upstream from the RNA 3' end; (ii) RL, GL and DNA:DNA pairing ten nucleotides upstream of the RNA 3' end are dispensable for anti-backtracking activity of NusG; (iii) LL may be involved in the anti-backtracking action of NusG, but is not critically important therein.

Mapping the effects of NusG on the TEC by DNA:DNA photocrosslinking with 8-methoxypsoralen

To directly test the effect of NusG on the reannealing of the upstream DNA, we developed a system to probe the DNA reannealing by photocrosslinking with 8-methoxypsoralen (8-MP). 8-MP specifically intercalates into the double-stranded 5'-TA-3' sequence and introduces a T-T inter-strand crosslink upon illumination with UV light (**Figure 4—figure supplement 1**). We designed a fully complementary TEC containing a unique 5'-TA-3' sequence motif positioned nine nucleotides upstream of the RNA 3' end (**Figure 4A, Figure 2—figure supplement 1**). The template DNA and the RNA primer were 5' labeled with ATTO680 to monitor DNA:DNA crosslinking and RNA extension by the denaturing PAGE. The system allowed us to probe DNA:DNA base pairing nine (TEC16), ten (TEC17) and eleven (TEC18) nucleotides upstream of the RNA 3' end by the stepwise extension of the RNA with subsets of NTPs (**Figure 4A**). The assembled TEC16 preparations produced only minute amounts of DNA:DNA crosslinked species as expected: the template DNA thymidine of TA site was anticipated to form the upstream-last base pair of the RNA:DNA hybrid. Upon formation of TEC17, the efficiency of DNA:DNA crosslinking remained low (<15%), despite the entire TA site being potentially available for pairing. In contrast, the crosslinking efficiency exceeded 40% in TEC18, comparable to that observed in a protein-free DNA:DNA duplex (~60%). TGT reduced crosslinking efficiency in TEC18 at least twofold (**Figure 4A**) consistent with its ability to stabilize TECs in the pre-translocated state (**Malinen et al., 2012**).

NusG effects on DNA:DNA crosslinking in the wild-type TEC18 were within the margins of the experimental errors (**Figure 4B**). However, NusG restored otherwise markedly reduced DNA crosslinking in Δ RL and Δ LL TEC18s. These observations suggest that NusG stabilizes the upstream DNA duplex 11 nucleotides upstream the RNA 3' end, but does not affect crosslinking in the wild-type TEC18 because the TA site is double-stranded even in the absence of NusG. Interestingly, NusG marginally but measurably enhanced crosslinking in the wild-type and Δ LL TEC17s, suggesting that NusG may affect the DNA conformation immediately upstream of the RNA:DNA hybrid (ten nucleotides upstream of the RNA 3' end).

Mapping the effects of NusG on the TEC by RNA:DNA photocrosslinking with 6-thioguanine

In light of the effect of NusG on the upstream DNA reannealing, it was reasonable to test its effect on a related and spatially adjacent process of RNA:DNA separation. We used template DNA-RNA photo-crosslinking by a guanine analogue 6-thioguanine (6-TG), to probe the accessibility of RNA to DNA in the absence and presence of NusG. The TECs contained the fully complementary transcription bubble and the 5' ATTO680-labeled 16-nucleotide RNA primer with the nine nucleotide annealing region. The initial TEC16 contained 6-TG in the template DNA eight base pairs upstream of the RNA 3' end (**Figure 5A, Figure 2—figure supplement 1**). Upon exposure to UV light, the TEC16

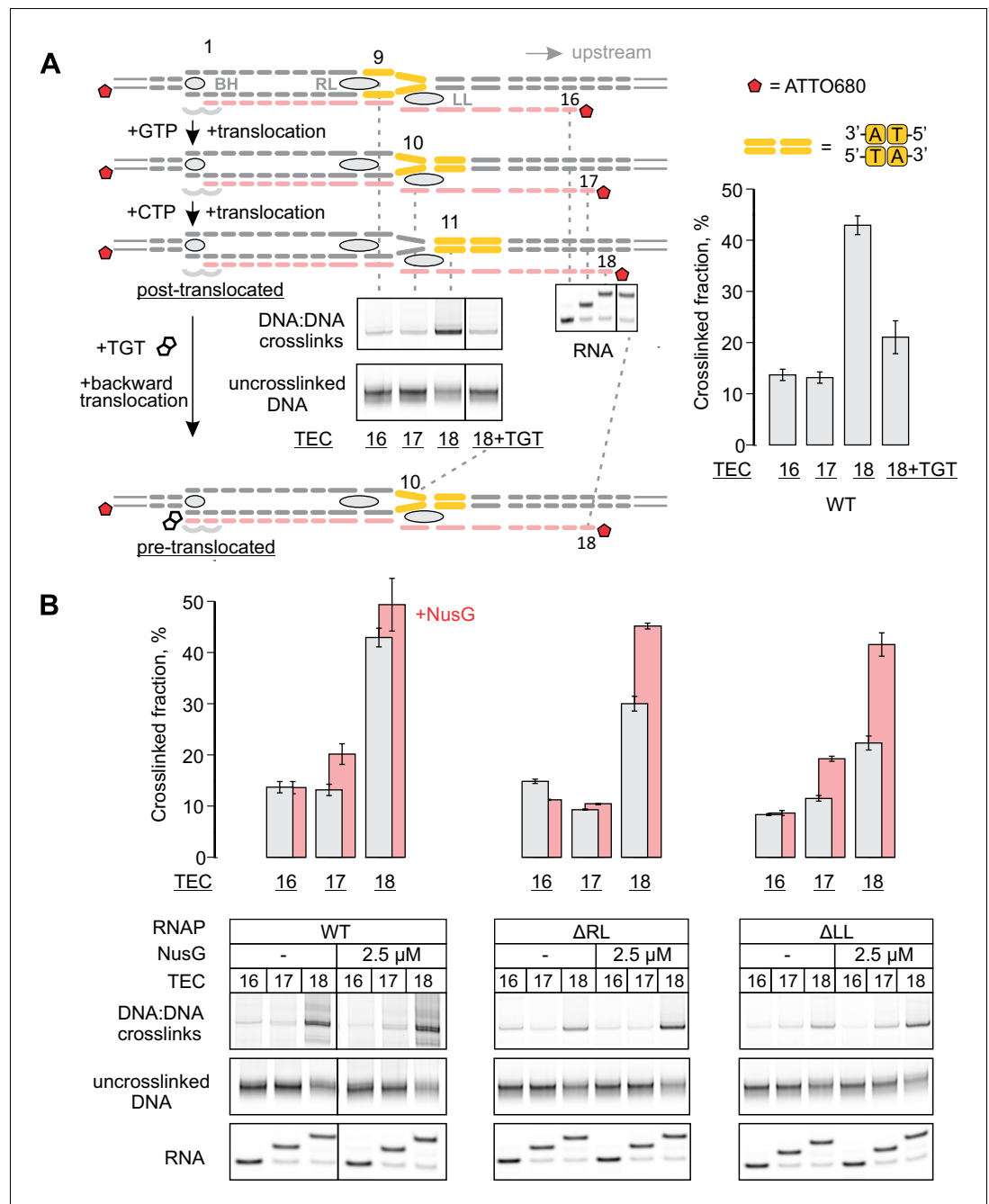


Figure 4. Probing the effects of NusG and deletions of the RNAP domains on the structure of the upstream fork junction by DNA:DNA photocrosslinking with 8-methoxypsoralen (8-MP). **(A)** TECs containing the unique 8-MP intercalation site were supplemented with 8-MP, walked by up to three nucleotides, supplemented with 5 μM TGT (where indicated) and illuminated with the UV light. **(B)** DNA:DNA crosslinking in TECs formed by the wild-type and altered RNAPs in the absence (grey bars) and presence (red bars) of NusG. Error bars indicate the range of duplicate measurements or SDs of several measurements (**Table 5**). The gel panels were spliced from the same gel and the pixel counts were linearly scaled to span the full 8 bit grayscale range within each panel. Joined panels have the same scaling.

DOI: [10.7554/eLife.18096.011](https://doi.org/10.7554/eLife.18096.011)

The following figure supplement is available for figure 4:

Figure supplement 1. Control experiment: the specificity of DNA:DNA photocrosslinking with 8-MP.

DOI: [10.7554/eLife.18096.012](https://doi.org/10.7554/eLife.18096.012)

produced crosslinked DNA-RNA species that migrated considerably slower than the un-crosslinked RNA primer in a denaturing gel (**Figure 5A**). Two major crosslinked species (a band and a smear) were observed that likely originated from 6-TG crosslinks to different RNA bases within the crosslinking range. Walking the RNAP along the DNA revealed that 6-TG efficiently crosslinks to the RNA primer eight (initial TEC16) and nine (TEC17) nucleotides upstream of the RNA 3' end. Crosslinking was largely abolished when the separation was increased to ten nucleotides in TEC18, yet restored when TEC18 was stabilized in the pre-translocated state by TGT (**Figure 5A**). Qualitatively similar results were obtained with the initial TEC16 containing 6-TG nine nucleotides upstream of the RNA 3' end (**Figure 5—figure supplement 1**); in this system, crosslinking was abolished upon an extension to form TEC17.

Deletion of the RL increased the overall intensity of crosslinks in TEC16 and TEC17, presumably by eliminating the protein domain that competed with RNA for crosslinking. More importantly, DNA efficiently crosslinked to RNA in Δ LL TEC18, but not in the Δ RL or the wild-type TEC18 (**Figure 5B**). These observations suggest that Δ LL makes the RNA accessible to DNA ten nucleotides upstream of the RNA 3' end, but do not necessarily suggest that the RNA:DNA hybrid is longer in Δ LL TEC. An RNA with a mismatch against 6-TG efficiently crosslinked to DNA in the wild-type TEC17, demonstrating that the crosslinks reflect the physical accessibility of RNA to DNA and do not require the RNA:DNA base pairing (**Figure 5—figure supplement 1**). NusG did not measurably affect the crosslinking efficiency eight, nine or ten nucleotides upstream of the RNA 3' end in wild-type and altered TECs (**Figure 5B**), indicating that NusG does not alter the accessibility of RNA to DNA at the upstream edge of the transcription bubble.

Mapping the effects of NusG on the TEC by a fluorescent beacon in the template DNA

We have previously reported that the base analogue fluorophore 6-MI positioned in template DNA strand within the 5'-TXG-3' beacon sequence (where X is 6-MI and G is a guanine functioning as a quencher) displays 2–5 fold brighter fluorescence when positioned nine and ten nucleotides upstream of the RNA 3' end relative to fluorescence levels observed eight and eleven nucleotides upstream of the RNA 3' end (**Figure 6A**). This system was originally designed for the time-resolved studies of the RNAP translocation (**Malinen et al., 2012**). Here we revisit this setup to complement the photocrosslinking techniques in assessing the effects of NusG on base pairing immediately upstream of the RNA:DNA hybrid and on the conformation of the upstream DNA.

The TECs contained the fully complementary transcription bubble and 16-nucleotide RNA primer with nine nucleotides annealing region (**Figure 6A, Figure 2—figure supplement 1**). The initial TEC16 contained 6-MI base in template DNA positioned eight base pairs upstream of the RNA 3' end. The TEC was walked along the DNA by up to three positions by the addition of subsets of NTPs and the 6-MI fluorescence was monitored (**Figure 6A**). We attribute the low fluorescence of 6-MI eight nucleotides upstream of the RNA 3' end (the assembled TEC16) to the quenching effect of the upstream guanine that forms the upstream-most base pair of the RNA:DNA hybrid. The elevated 6-MI fluorescence nine nucleotides upstream of the RNA 3' end is likely due to the unstacking of the upstream quenching guanine because it no longer belongs to the RNA:DNA hybrid. The elevated fluorescence persists ten nucleotides upstream of the RNA 3' end, arguing that 6-MI does not reestablish the stacking interaction with the upstream guanine. Finally, the 6-MI fluorescence is reduced 11 nucleotides upstream from the RNA 3' end, likely because 6-MI reestablishes the stacking interaction with the upstream guanine at this position. An experiment performed in the presence of TGT suggests that increased 6-MI fluorescence at the upstream edge of the RNA:DNA hybrid originates from a post-translocated TEC (**Figure 2D**).

The major effect of NusG on the wild-type TEC was the increase in the fluorescence of 6-MI at the upstream edge of the RNA:DNA hybrid (in TEC17). The effect was also observed in Δ LL and Δ RL TEC17s (**Figure 6C, Figure 6—figure supplement 1**). The rate of fluorescence increase in the wild-type TEC17 was dependent on NusG concentration with $K_d \sim 120$ nM (**Figure 6C**). NusG binds about 30 Å from 6-MI at the upstream edge of the RNA:DNA hybrid and is therefore unlikely to affect the 6-MI fluorescence directly. Instead, we suggest that NusG increases 6-MI fluorescence by repositioning the quenching guanine immediately upstream of the RNA:DNA hybrid.

In the absence of LL, NusG significantly changed 6-MI fluorescence also in TEC18 and TEC19 (**Figure 6B**). Remarkably, Δ LL TECs deviated most from the wild-type TECs in the absence of NusG

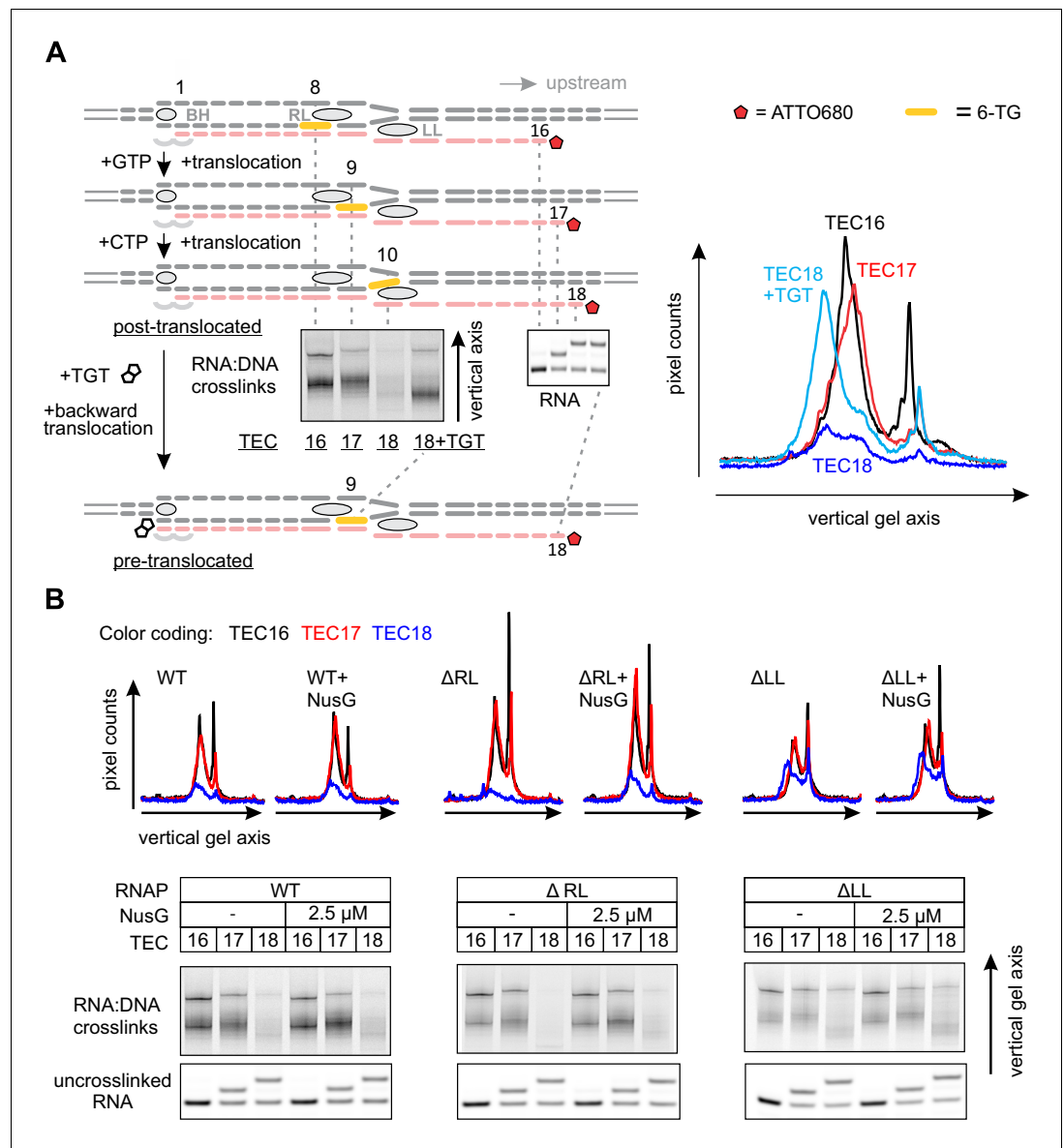


Figure 5. Probing the effects of NusG and deletions of the RNAP domains on the structure of the upstream fork junction by RNA:DNA photocrosslinking with 6-TG. (A) TECs containing 6-TG in the template DNA were walked by up to three nucleotides, supplemented with 5 μM TGT (where indicated) and illuminated with the UV light. (B) RNA:DNA crosslinking in TECs formed by the wild-type and altered RNAPs in the absence and presence of NusG. The gel panels were spliced from the same gel and the pixel counts were linearly scaled to span the full 8 bit grayscale range within each panel. The pixel intensity profiles for each gel lane are shown above the gels. The independent repeats are presented in [Figure 5—figure supplement 2](#). DOI: [10.7554/eLife.18096.013](https://doi.org/10.7554/eLife.18096.013)

The following figure supplements are available for figure 5:

Figure supplement 1. RNA:DNA photocrosslinking with 6-TG reflects the physical accessibility of RNA to DNA. DOI: [10.7554/eLife.18096.014](https://doi.org/10.7554/eLife.18096.014)

Figure supplement 2. The effects of NusG (A), TGT (B) and ΔLL (C) on RNA:DNA photocrosslinking with 6-TG. DOI: [10.7554/eLife.18096.015](https://doi.org/10.7554/eLife.18096.015)

but displayed the identical fluorescence intensities in the presence of NusG (**Figure 6B**). Overall, the effects of NusG on 6-MI fluorescence largely paralleled its effects on the DNA crosslinking with 8-MP leading to similar conclusions: (i) NusG likely affects the DNA conformation immediately

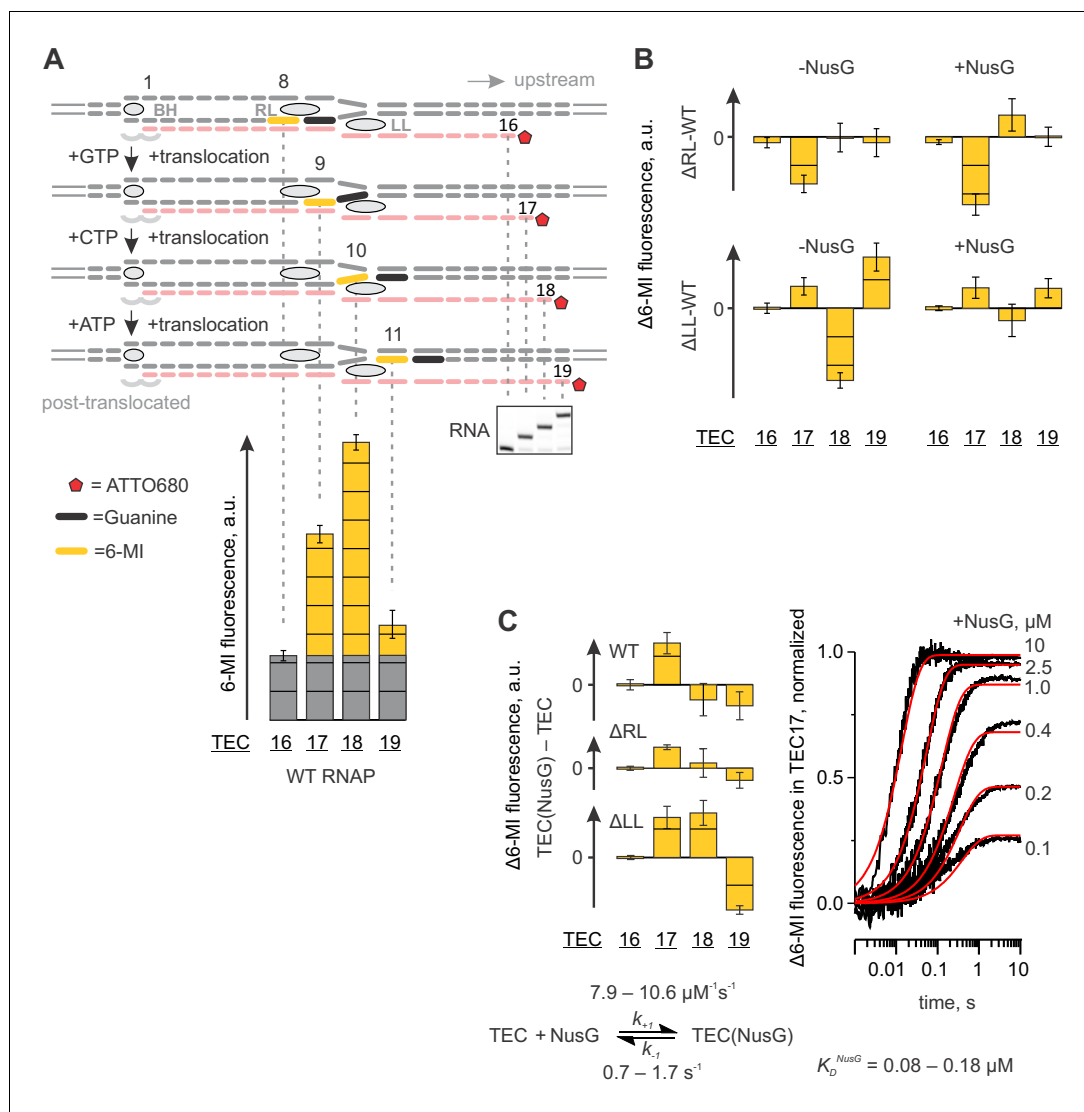


Figure 6. Probing the effects of NusG and deletions of the RNAP domains on the structure of the upstream fork junction by a fluorescent beacon in the template DNA. Error bars indicate the range of duplicate measurements or SDs of several measurements (Table 5). (A) Walking 6-MI (yellow dash) through the upstream fork junction modulates 6-MI stacking with the upstream guanine (black dash) that functions as a strong quencher. (B) The effects of ΔRL and ΔLL on the TEC fluorescence in the absence and the presence of NusG. (C) The effects of NusG on 6-MI fluorescence of the wild-type and altered TECs. Monitoring the fluorescence of TEC17 upon mixing with NusG in a stopped flow instrument (graph on the right, black curves) allows for the estimation of the binding and the dissociation rate constants. The analysis scheme is depicted below the graph. The best fit curves (red) were simulated using $k_{+1}=9.2 \mu M^{-1}s^{-1}$; $k_{-1}=1.1 s^{-1}$. The lower and upper bounds of rate constants were calculated by combined analysis of data from two independent experiments by FitSpace routine of Kintek Explorer software (at a 10% increase in Chi2).

DOI: 10.7554/eLife.18096.016

The following figure supplement is available for figure 6:

Figure supplement 1. Primary data for Figure 6: the effects of NusG and deletions of the RNAP domains on 6-MI fluorescence.

DOI: 10.7554/eLife.18096.017

upstream of the RNA:DNA hybrid and (ii) NusG reverses the alterations in the upstream DNA conformation introduced by deletion of the LL. However, in contrast to 8-MP crosslinking experiments,

NusG did not compensate for and, instead, increased the differences in the fluorescence intensities between the wild-type and Δ RL TECs (**Figure 6B**).

An overview: NusG and the structure of the upstream fork junction

NusG likely inhibits backtracking by acting on the upstream DNA that, at the time of writing, is universally absent from the crystal structures of bacterial TECs. At the same time, the conformation of the upstream DNA in published TEC models (*Opalka et al., 2010; Martinez-Rucobo et al., 2011; Andrecka et al., 2009*) as well as recent X-ray (*Barnes et al., 2015*) and CryoEM (*Bernecky et al., 2016*) structures of RNA polymerase II are incompatible with the structure of bacterial TEC (in case of RNA polymerase II models), the upstream DNA mapping data presented here (see below), and/or NusG binding. Similarly, the conformation of the upstream DNA resolved in the crystal structures of the bacterial initiation complexes (*Zuo and Steitz, 2015; Bae et al., 2015; Liu et al., 2016*) is strongly influenced by the sigma factor and therefore is not suitable for modeling of the upstream fork junction in the TEC. To gain the mechanistic insights into the anti-backtracking action of NusG, we used our data to generate an accurate map of the upstream fork junction (**Figure 7B**) and further employed it to postulate a tentative structural model of a NusG-TEC complex with the upstream DNA (**Figures 1, 8**).

Together, the 8-MP DNA:DNA photocrosslinking (\pm TGT), 6-TG DNA:RNA photocrosslinking (\pm TGT), 6-MI fluorescence pattern and crystal structures of bacterial TECs lacking the upstream DNA (*Vassilyev et al., 2007*) define the resting TECs as (i) post-translocated, (ii) containing nine base pairs RNA:DNA hybrid, and (iii) containing the upstream DNA duplex that starts eleven nucleotides upstream of the RNA 3' end. The experiments with TECs containing mismatched non-template DNA strand additionally suggest that the upstream DNA base pairs ten nucleotides upstream of the RNA 3' end (**Figure 7A**). Specifically, the 8-MP crosslinks the TA site eleven nucleotides upstream of the RNA 3' end only when DNA is matched ten nucleotides upstream of the RNA 3' end (**Figure 7A, Figure 7—figure supplement 2**). Similarly, TECs containing a DNA:DNA mismatch either directly against or one nucleotide upstream of 6-MI start to differ in fluorescence levels from the matched TECs when the mismatch is ten nucleotides from the RNA 3' end (**Figure 7A, Figure 7—figure supplement 1**). Therefore, DNA likely base pairs ten nucleotides from the RNA 3' end, but the first upstream base pair deviates from the geometry of the conventional B-form DNA duplex and/or is highly dynamic (**Figure 8B**).

To generate a NusG-TEC model, we positioned NusG NTD over the β' clamp helices in bacterial TEC lacking the upstream DNA (*Vassilyev et al., 2007*) guided by the model of NusG-RNAP complex in (*Martinez-Rucobo et al., 2011*). We then modelled in the upstream duplex DNA following the overall direction suggested by the published models (*Opalka et al., 2010; Andrecka et al., 2009; Bernecky et al., 2016*) but avoiding clashes with the NusG NTD and maintaining a canonical B-duplex as far downstream as possible. Joining the upstream duplex with the RNA:DNA hybrid required altering the sugar-phosphate backbone of the downstream-most nucleotides of the upstream DNA but allowed maintaining the DNA base pairing. In the resulting model (**Figure 1, 8**), the upstream DNA has an ample space to move away from the NusG NTD and the β' clamp surfaces towards the cleft between the β_1 and β flap domains by hinging around the downstream most base pair (position ten in **Figure 8**), reflecting the natural flexibility of the upstream DNA (*Coban et al., 2006*). However, we suggest that the conformation where the upstream DNA lines β' clamp and NusG NTD (**Figure 1, 8**) is the most relevant to the NusG effects on transcription elongation. In such a scenario, the two downstream-most base pairs of the upstream DNA (positions ten and eleven in **Figure 8**) occupy a narrow channel walled by NusG NTD and LL, thereby explaining the functional interactions between the NusG, the LL and the upstream DNA in RNA cleavage, cross-linking and fluorescence assays.

Discussion

NusG inhibits backtracking by acting on the upstream DNA

E. coli NusG was shown to enhance elongation in vitro over two decades ago, acting mainly by reducing pausing (*Burova et al., 1995*). More specifically, NusG was found to reduce backtracked but not hairpin-stimulated pauses (*Artsimovitch and Landick, 2000; Pasman and von Hippel,*

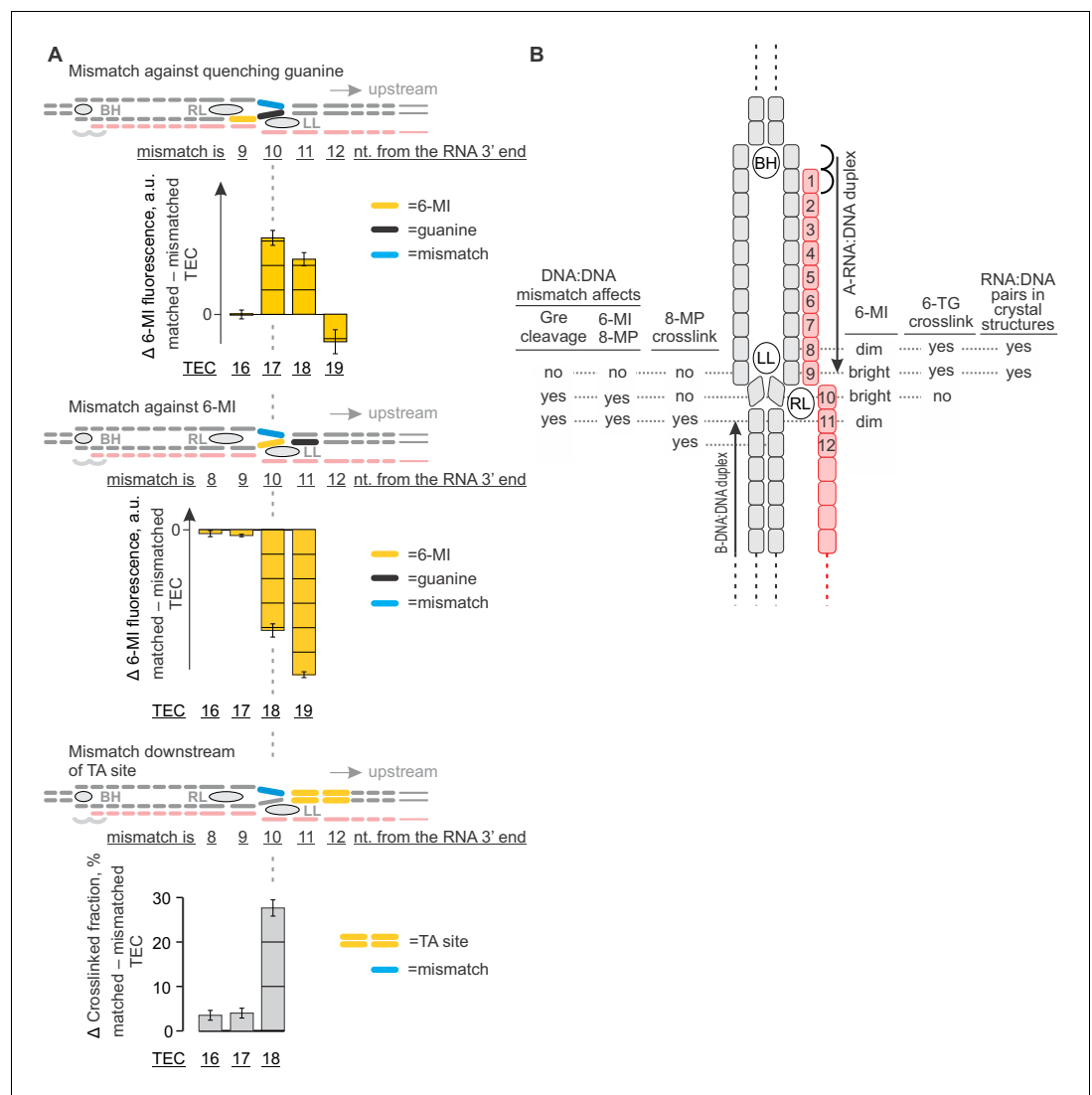


Figure 7. The effects of DNA mismatches suggest the minimal transcription bubble. (A) DNA:DNA mismatches against quenching guanine (top) and 6-MI (middle), or downstream of the TA site (bottom) alter the TEC properties when positioned ten nucleotides upstream of the RNA 3' end and further upstream. Error bars indicate the range of duplicate measurements or SDs of several measurements (Table 5). (B) Mapping the upstream edge of the transcription bubble based on data in Figures 3C, 4A, 5A, 6A and 7A.

DOI: 10.7554/eLife.18096.018

The following figure supplements are available for figure 7:

Figure supplement 1. Primary data for Figure 7A: the effects of DNA mismatches on 6-MI fluorescence.

DOI: 10.7554/eLife.18096.019

Figure supplement 2. Primary data for Figure 7A: the effect of a DNA mismatch downstream of the TA site on crosslinking with 8MP.

DOI: 10.7554/eLife.18096.020

2000). It was later established that NusG-family proteins bind to, and bridge, the β' clamp with the β lobe across the RNAP cleft (Belogurov et al., 2007; Mooney et al., 2009b; Klein et al., 2011; Martinez-Rucobo et al., 2011). Biochemical studies further concluded that the archaeal NusG orthologue Spt5 (Hirtreiter et al., 2010) and the specialized NusG paralogue RfaH (Sevostyanova et al., 2011) enhance transcription elongation by stabilizing the β' clamp in a closed conformation. However, *E. coli* RfaH accelerates RNAP at pause sites known to involve clamp opening, as well as at the backtracked pauses, whereas *E. coli* NusG has only marginal effect at the former sites (Kolb et al.,

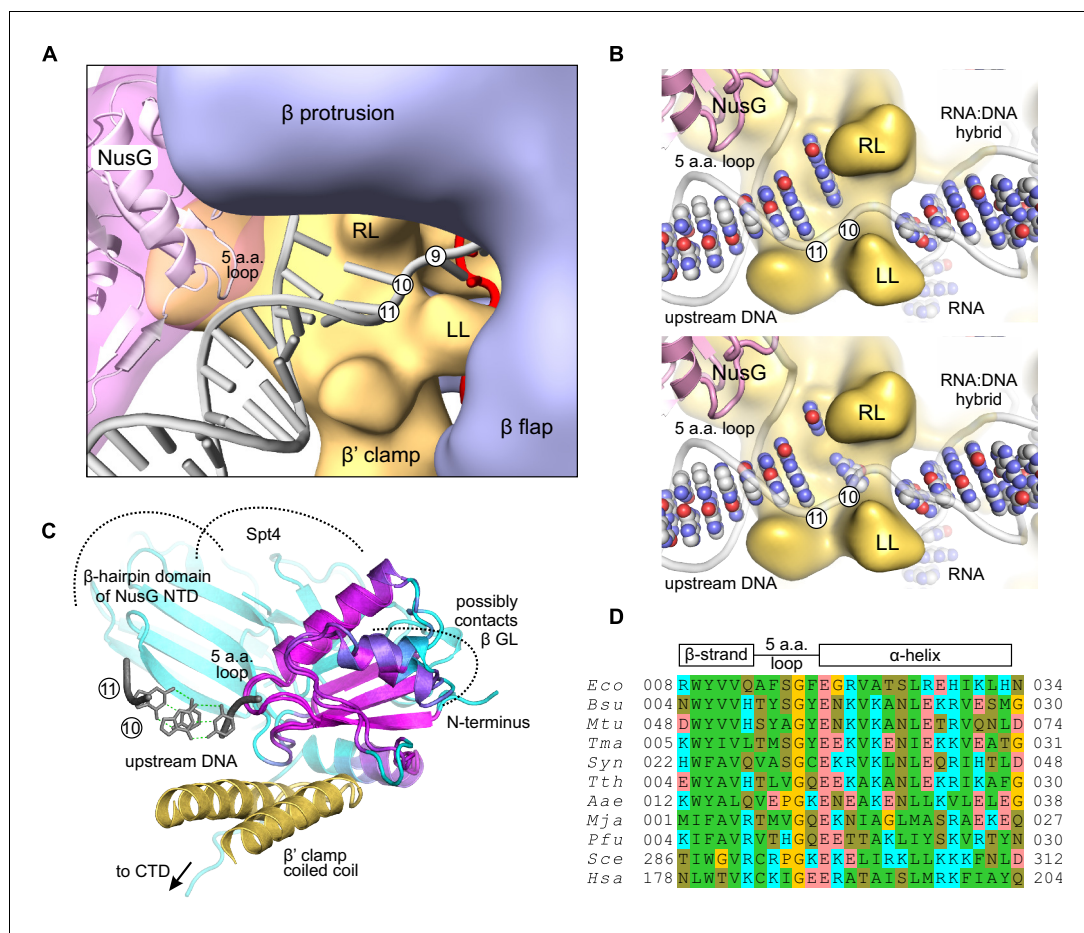


Figure 8. A model of the upstream fork junction. DNA bases are numbered from the RNA 3' end in the post-translocated TEC. (A) An overview: β' LL and the structurally conserved five amino acid loop of NusG NTD form a channel accommodating the exiting upstream DNA. (B) The template DNA nucleotide at position ten can be modeled to pair with the non-template DNA in a partially unstacked conformation (top) or interact with the cleft between the β' RL and β' LL (bottom). DNA and RNA bases are shown as spheres, sugar-phosphate backbones as cartoons. The β subunit is omitted for clarity. (C) The superimposition of *Aquifex aeolicus* NusG NTD (PDB ID 1M1G) and *Pyrococcus furiosus* Spt5 NTD (PDB ID 3P8B). The cartoon color changes from magenta in well superimposed regions to cyan in poorly superimposed or unaligned regions. The β' clamp helices (yellow, from PDB ID 3QQC), the upstream DNA (grey, from the model in A) and Spt4 (cyan cartoon, from PDB ID 3P8B) are included to present the superimposition in the context of the TEC. (D) Multiple sequence alignment of the structurally conserved five amino acid loop of NusG family proteins and the flanking secondary structure elements. Species names are abbreviated as follows: *Eco*, *E. coli*, *Bsu*, *Bacillus subtilis*, *Mtu*, *Mycobacterium tuberculosis*, *Tma*, *Thermotoga maritima*, *Syn*, *Synechocystis sp. PCC 6803*, *Tth*, *T. thermophilus*, *Aae*, *Aquifex aeolicus*, *Mja* *Methanocaldococcus jannaschii*, *Pfu*, *Pyrococcus furiosus*, *Sce*, *Saccharomyces cerevisiae*, *Hsa*, *Homo sapiens*. Amino acid residues are shaded as follows: hydrophobic –green, polar –olive, Pro and Gly –yellow, Asp and Glu –red, Arg, Lys and His –cyan.

DOI: [10.7554/eLife.18096.021](https://doi.org/10.7554/eLife.18096.021)

The following source data is available for figure 8:

Source data 1. NusG-TEC model.

DOI: [10.7554/eLife.18096.022](https://doi.org/10.7554/eLife.18096.022)

2014; Anthony et al., 2000; Belogurov et al., 2010). Moreover, crystal structures of the backtracked TECs revealed the closed clamp (Sekine et al., 2015; Wang et al., 2009), whereas the specificity of *E. coli* NusG for backtracked pauses was reaffirmed in single molecule experiments (Herbert et al., 2010). Together, these observations suggest that *E. coli* NusG enhances transcription elongation by means other than restricting the β' clamp movement.

Here, we show that NusG slows backtracking but does not affect the on-pathway elongation in the non-paused TEC used in our study. In contrast, Herbert *et al* (Herbert *et al.*, 2010) concluded that NusG has a modest stimulatory effect (10–20%) on the pause-free elongation rate, in addition to inhibiting backtracking. One possibility is that a subset of TECs backtracked by one nucleotide display the elongation rates within the pause-free range compiled by Herbert *et al.* Alternatively, NusG may have a marginal effect on the elongation rate in a subset of the non-paused on-pathway TECs with yet to be identified sequence determinants. In any case, the effect of NusG on the pause-free elongation rate estimated by Herbert *et al* is small comparing with the specific effect of NusG on the backtracking rate (~2.5 fold) that we report here.

We further demonstrate that two DNA mismatches immediately upstream of the RNA:DNA hybrid increase the backtracking rate and render the TEC insensitive to NusG. In contrast, NusG reduced backtracking normally in a TEC missing the GL, a contact point with the β subunit (Figure 1) that is required for anti-pausing by *E. coli* RfaH (Sevostyanova *et al.*, 2011). β' RL was similarly dispensable, whereas the β' LL was slightly stimulatory for the anti-backtracking activity of NusG. Overall, our data suggest that the intrinsic action of NusG on the *E. coli* TEC is restricted to inhibiting backtracking and is exclusively mediated through the upstream fork junction. Remarkably, several other transcription factors, including Mfd (Deaconescu *et al.*, 2006) and UvrD (Epshtein *et al.*, 2014) that link transcription to DNA repair operate through the upstream fork junction.

Interestingly, Δ GL TEC backtracked two-fold slower than the wild-type TEC both in the presence and in the absence of NusG (Figure 3C). GL is located more than 20Å from the duplex DNAs and the RNA:DNA hybrid but may directly contact the single-stranded non-template DNA. Accordingly, we suggest that GL promotes backtracking by altering the conformation of the non-template DNA in a manner that increases the propensity of the TEC to backtrack, e.g. by facilitating the downstream DNA re-annealing. Indeed, GL restricts the downstream portion of the single-stranded non-template DNA within the main channel in the initiation complex (Zhang *et al.*, 2012) and may therefore have a similar functionality in the TEC.

Mapping the structure of the upstream fork junction

The upstream DNA decisively emerged as the major determinant of NusG anti-backtracking effect. However, the upstream DNA is absent from the crystal structures of bacterial TECs, the published models are incompatible with NusG binding, and even the register of the upstream DNA reannealing is uncertain. To gain mechanistic insights into the anti-backtracking action of NusG, we performed a comprehensive mapping of the upstream fork junction using fluorescent base analogues and site-specific crosslinking. Our data are fully consistent with the nine base pairs RNA:DNA hybrid and the upstream DNA duplex that starts eleven base pairs upstream of the RNA 3' end in the post-translocated TEC (Figure 7B). The effects of DNA mismatches additionally suggest that DNA is paired immediately upstream of the RNA:DNA hybrid. In combination, the data obtained with the matched and mismatched TECs suggest that the first pair of the upstream DNA is unstacked from both the RNA:DNA hybrid and the rest of the upstream DNA.

We then combined our mapping data with the model of the NusG-RNAP complex (Martinez-Rucobo *et al.*, 2011) to sketch a NusG-TEC model with the upstream DNA. We found that in a subset of spatially feasible conformations of the upstream DNA (see results), the two downstream-most base pairs are accommodated in a narrow 'exit' channel walled by the NusG and the LL (Figure 1, 8). Such an arrangement plausibly explains the cooperation between NusG and the LL in stabilizing the upstream DNA pairing and inhibiting backtracking. Notably, NusG loop facing the upstream DNA is strictly conserved in size in bacteria and archaea (five amino acids, Figure 8), but the evidence for the conserved residue-specific contacts between the NusG and the upstream DNA is lacking. We propose that NusG provides a complementary molecular surface to the paired upstream DNA and possibly also affects the overall direction of the upstream DNA duplex. Finally, NusG likely stabilizes DNA pairing only in a subset of the upstream DNA conformations, yet influences the overall backtracking rate by targeting those conformations that are most favorable for backtracking, *i.e.*, the non-template DNA is optimally positioned for the strand exchange with the RNA.

NusG inhibits backtracking by stabilizing the minimal bubble in the pre-translocated TEC

NusG inhibits backtracking by stabilizing the DNA base pair eleven nucleotides upstream of the RNA 3' end (**Figure 3C**). This pair corresponds to the second DNA pair upstream of the RNA:DNA hybrid in the post-translocated TEC (pair eleven in **Figure 8**). However, NusG does not affect the equilibrium between the pre- and post-translocated states (**Figure 2C–D**) and therefore likely acts on the pre-translocated TEC where the base pair eleven nucleotides upstream of the RNA 3' end lies immediately upstream of the RNA:DNA hybrid (pair ten in **Figure 8**, see also schematics in **Figure 3C**). Accordingly, we propose that NusG facilitates DNA pairing immediately upstream of the RNA:DNA hybrid in the pre-translocated TEC, thereby reducing backtracking.

It remains uncertain how NusG inhibits backtracking without affecting the equilibrium between the post- and pre-translocated states (**Figure 2C–D**). One possible explanation is that backtracking and backward translocation are limited by different processes in our system. It has been hypothesized that, rather than moving in sync along the different nucleic acid chains of the TEC, RNAP moves forward by sequentially translocating the RNA:DNA hybrid and the downstream DNA (**Brueckner and Cramer, 2008**). The synchronous sliding of RNAP along the nucleic acids is only superficially plausible in the 2D schematics (**Figures 2–7**) but is much less likely in the actual 3D structure of the TEC (**Figure 1**) where the downstream DNA and the RNA:DNA hybrid are separated by a 90° bend. Analogously, the backward translocation may involve the sequential translocation of the downstream DNA and the RNA:DNA hybrid. The former process is NusG independent and may limit the rate of the backward translocation, whereas the latter process is modulated by NusG and may be thus rate limiting for backtracking. We further argue that the difficulty of reconciling the large effect of NusG on backtracking with its small (**Herbert et al., 2010**) or undetectable (this work) effect on the forward and backward translocation in the context of a single-step translocation model lends support to a two-step translocation mechanism.

Three distinct effects of NusG NTD on the TEC

The available data suggest three independent and structurally plausible effects of NusG NTD on the TEC. First, NusG binds near the upstream fork junction and stabilizes the upstream DNA duplex, thereby inhibiting spontaneous backtracking at most template positions ([**Herbert et al., 2010**] and this work). Second, the NTD restricts the conformational flexibility of the β' clamp, with different outcomes for the transcription elongation. Archaeal Spt5 (**Hirtreiter et al., 2010; Schulz et al., 2016**) and *E. coli* RfaH (**Sevostyanova et al., 2011**) exert at least part of their elongation enhancing effects by stabilizing the clamp. In contrast, clamp stabilization by *E. coli* NusG only marginally contributes to anti-pausing (**Kolb et al., 2014; Belogurov et al., 2010**). Third, NusG binds to specific sequences in the single-stranded non-template DNA, thereby introducing infrequent yet physiologically relevant pauses (**Yakhnin et al., 2016, 2008**) and facilitating intrinsic termination in some species (**Czyz et al., 2014**). Similar effects are well documented for other dissociable factors and TEC components positioned near the single stranded non-template DNA (**Perdue and Roberts, 2011; Artsimovitch and Landick, 2002; Vvedenskaya et al., 2014; Arimbasseri and Maraia, 2015**). Interestingly, *T. thermophilus* NusG slows down the cognate RNAPs at non-paused sites by an unknown mechanism (**Sevostyanova and Artsimovitch, 2010**). It remains to be determined if this unusual effect is mediated through the clamp or contacts with the non-template and upstream DNA.

Functional significance of the intrinsic stimulation of transcription elongation by NusG family proteins

We argue that anti-backtracking represents the only conserved functionality of NusG family proteins. Backtracking is a universally conserved and functionally important feature of the multisubunit RNAPs that has been documented in vitro and in vivo in both bacteria and eukaryotes (reviewed in [**Nudler, 2012**]). The stimulation of RNA chain elongation by NusG has been documented in vitro for bacterial (**Burova et al., 1995**), archaeal (**Hirtreiter et al., 2010**) and eukaryotic (**Wada et al., 1998**) transcription systems. The NTD is sufficient for these effects on elongation, but the structural elements that superimpose well between the bacterial NusG NTD and the archaeal Spt5 NTD (**Figure 8C**) are limited to (i) the beta sheet that comprises the RNAP-binding site, (ii) the conserved five amino acid loop that we implicate in the anti-backtracking action of *E. coli* NusG, and (iii) the

N-terminus of the α -helix that follows this loop and possibly interacts with the single-stranded non-template DNA (Crickard *et al.*, 2016). The lack of the strong conservation of the surface residues (Figure 8D) suggests that the anti-backtracking activity may be determined by the overall fold of the NusG NTD and is only weakly dependent on the nature of the individual amino acid side-chains, consistent with the mutational analysis of *E. coli* NusG (Mooney *et al.*, 2009b).

While the above considerations suggest that the stimulation of transcription elongation by NusG family proteins may be important for long-term survival and fitness, in the experimental systems studied to date, functional contacts established by the CTD, such as Rho and S10 in *Bacteria*, appear to be more critical. In *E. coli*, the essential function of NusG is to facilitate termination of transcription by Rho, thereby maintaining the operon borders (Cardinale *et al.*, 2008), suppressing pervasive antisense transcription (Peters *et al.*, 2012), and inhibiting R-loop formation (Krishna Leela *et al.*, 2013). Stimulating Rho-dependent termination is also likely the major, albeit a non-essential, function of *B. subtilis* NusG (Ingham and Furneaux, 2000). NusG and its paralog RfaH have also been proposed to mediate transcription-translation coupling via direct contacts with S10 (Burmam *et al.*, 2010, Burmam and Rösch, 2011). However, the CTD contacts are not universally conserved: Rho is absent in eukaryotes and even some *Bacteria* (D'Heygère *et al.*, 2013), whereas S10 and RNAP are separated by a nuclear membrane in eukaryotes, where CTD interacts with proteins involved in splicing, polyadenylation, and other RNA processing pathways.

In eukaryotes, the intrinsic stimulatory activity of Spt5 NTD on transcription elongation is non-essential, but abolishing it leads to the temperature sensitive phenotypes (Crickard *et al.*, 2016). In rare circumstances, the stimulatory effect of Spt4/5 may possibly be deleterious: Stp4/5 has been suggested to facilitate transcription through toxic repeat sequences in eukaryotes, thereby contributing to the progress of neurodegenerative disorders (Kramer *et al.*, 2016). Overall, we suggest that the universal conservation of the intrinsic stimulatory activity of NusG family proteins on transcription elongation underscores its importance, but the quantitative assessment of the *in vivo* role of this functionality in bacteria necessitates the analysis of transcription systems that natively lack Rho and Gre factors, e.g., those of Cyanobacteria.

Materials and methods

Reagents and oligonucleotides

DNA and RNA oligonucleotides were purchased from IBA Biotech (Göttingen, Germany) and Fidelity Systems (Gaithersburg, MD, USA). DNA oligonucleotides and RNA primers are listed in Table 2. NTPs were from Jena Bioscience (Jena, Germany). Tagetitoxin (TGT) was from Epicentre (Madison, WI, USA), 8-methoxypsoralen (8-MP) was from Sigma (St. Louis, MO, USA). The following buffers were used for the TEC assembly and transcription assays: TB0 (40 mM HEPES-KOH pH 7.5, 80 mM KCl, 5% glycerol, 0.1 mM EDTA, and 0.1 mM DTT), TB1 (TB0 supplemented with 1 mM MgCl₂), TB2 (TB0 supplemented with 2 mM MgCl₂ and 300 mM KCl) and TB10 (TB0 supplemented with 10 mM MgCl₂).

Proteins

All proteins were expressed in *E. coli* Xjb(DE3) (Zymo Research, Irvine, CA). The wild-type, Δ LL (β' Δ P251-S263 \rightarrow GG), Δ RL (β' Δ N309-K325) and Δ GL (β Δ R368-P376 \rightarrow GG) RNAPs were purified by Ni-, heparin and Q-sepharose chromatography as described previously (Svetlov and Artsimovitch, 2015). *E. coli* NusG was captured from the lysate by Ni-sepharose, the N-terminal hexa-histidine tag was cleaved by TEV-protease, imidazole was removed by dialysis, and the un-cleaved NusG, the cleaved tag and the TEV-protease were absorbed by passing the NusG solution over the Ni-sepharose. *E. coli* GreA containing C-terminal hexa-histidine tag was captured from lysate by Ni-sepharose followed by gel filtration as described (Perederina *et al.*, 2006). All proteins were dialyzed against the storage buffer (50% glycerol, 20 mM Tris-HCl pH 7.9, 150 mM NaCl (1M NaCl for GreA), 0.1 mM EDTA, 0.1 mM DTT) and stored at -20°C . Plasmids are listed in Table 3. Sequences of the plasmids are provided as Supplementary file 5 (plasmids.fas).

Table 2. DNA oligonucleotides and RNA primers used in this study.

Name	type	Sequence (5'→3')	Employment
S041M	tDNA	GCTACTCTACTGACATGATGCCTCCTCT $\underline{\text{X}}$ GAACTTAGATCGCTACAAGT	Figures 2,6–7
S154S	tDNA	GCTACTCTACTGACATGATGCCTCCTCTG $\underline{\text{S}}$ AACCTTAGATCGCTACAAGT	Figure 5—figure supplement 1
S155S	tDNA	GCTACTCTACTGACATGATGCCTCCTCT $\underline{\text{S}}$ GAACTTAGATCGCTACAAGT	Figure 5
S042	ntDNA	ACTTGTAGCGATCTAAGGTTCCAGAGGAGGCATCATGTCAGTAGAGTAGC	Figures 2,5–7
S150	ntDNA	ACTTGTAGCGATCTAAGGTT $\underline{\text{A}}$ CAGAGGAGGCATCATGTCAGTAGAGTAGC	Figure 7A
S056M	tDNA	GCTACTCTACTGCAATGACGCTCCTCT $\underline{\text{X}}$ GAACTTAGATCGCTACAAGT	Figures 2C, 3B, 7A
S076	tDNA	GCTACTCTACTGCAATGACGCTCCTCTGGAACCTTAGATCGCTACAAGT	Figure 3
S057	ntDNA	ACTTGTAGCGATCTAAGGTTCCAGAGGAGACGTCATTGCAGTAGAGTAGC	Figures 2C,3,7A
S152	ntDNA	ACTTGTAGCGATCTAAGGTT $\underline{\text{C}}$ GAGAGGAGACGTCATTGCAGTAGAGTAGC	Figures 3C,7A
S153	ntDNA	ACTTGTAGCGATCTAAGGTT $\underline{\text{G}}$ GAGAGGAGACGTCATTGCAGTAGAGTAGC	Figure 3C
S173	ntDNA	ACTTGTAGCGATCTAAGGTT $\underline{\text{AA}}$ AGAGGAGACGTCATTGCAGTAGAGTAGC	Figure 3—figure supplement 3
S114	tDNA	CGTACTCTACTCGAATAGCATCTCCTCTGGAACCTTAGATCGTCACAAGT	Figure 3C
S115	ntDNA	ACTTGTAGCGATCTAAGGTTCCAGAGGAGATGCTATTTCGAGTAGAGTAGC	Figure 3C
S170	tDNA	Atto680- TGGTGTCTGCTGTCCGCTGCCTCCTCTGTAGTCTGTGCTCGTGTCTGGT	Figures 4,7A
S171	ntDNA	ACCAGACACGAGCACAGACTACAGAGGAGGCAGACGGACAGCAGACACCA	Figures 4,7A
S224	ntDNA	ACCAGACACGAGCACAGACT $\underline{\text{AA}}$ AGAGGAGGCAGACGGACAGCAGACACCA	Figure 7A
R024	RNA	Atto680- CUCACAACCAGAGGAG	Figures 2,5–7
R052	RNA	Atto680- CUCACAACCAGAGGAG $\underline{\text{Y}}$ C	Figure 3
R079	RNA	Atto680- CAACACAACAGAGGAG	Figures 4,7, Figure 5—figure supplement 1

$\underline{\text{X}}$ = 6-methyl-isoxanthopterin; $\underline{\text{S}}$ = 6-thioguanine; $\underline{\text{Y}}$ = 2-aminopurine

Mismatches in ntDNA are marked in blue and underlined.

R079 is a chimeric oligo: six 5' nucleotides are DNA.

DOI: 10.7554/eLife.18096.023

TEC assembly

TECs (1 μM) were assembled by a procedure developed by Komissarova et al (Komissarova et al., 2003). An RNA primer was annealed to the template DNA, incubated with RNAP for 10 min, and with the non-template DNA for 20 min at 25°C. RNA, template DNA, non-template DNA and RNAP were present at 1–2 μM during assembly. The exact ratios between the TEC components in different assays are listed in Table 4. The assembly was carried out in TB0 buffer for TECs used in backtrack-ing, RNA cleavage and dinucleotide release experiment or in TB10 buffer for TECs used in all the other experiments. In the nucleotide addition experiments the assembled TEC16 were used. In the pyrophosphorolysis and NusG binding experiments the assembled TEC16s were pre-extended into TEC17s with 5 μM ATP or GTP, respectively. In the former case, the excess of ATP was further

Table 3. *E. coli* protein expression vectors used in this study.

Name	Description	Source/reference
pVS10	wild-type RNAP (T7p- α - β - β' -His ₆ -T7p- ω)	(Belogurov et al., 2007)
pTG011	$\Delta\beta\text{GL}$ RNAP (T7p- α -His ₆ - β [$\Delta\text{R368-P376}\rightarrow\text{GG}$]- β' - ω)	this work
pMT041	$\Delta\beta'$ RL RNAP (T7p- α - β - β' [$\Delta\text{N309-K325}$]-TEV_His ₁₀ -T7p- ω)	this work
pHM001	$\Delta\beta'$ LL RNAP (T7p- α - β - β' [$\Delta\text{P251-S263}\rightarrow\text{GG}$]-TEV_His ₁₀ -T7p- ω)	this work
pIA578	GreA (T7p-GreA_His ₆)	(Perederina et al., 2006)
pGB043	NusG (T7p-His ₆ -TEV_NusG)	made by GB in Artsimovitch lab.

Sequences of the plasmids are provided as Supplementary file 5 (plasmids.fas).

DOI: 10.7554/eLife.18096.024

removed by passing through the desalting spin columns (40K cutoff) with TB10 buffer. In the RNA cleavage experiments TEC18s were directly assembled in TBO buffer.

Nucleotide addition and RNA cleavage measurements

Time-resolved measurements of nucleotide addition were performed in an RQF 3 quench-flow instrument (KinTek Corporation, Austin, TX). The reaction was initiated by rapid mixing of 14 μ l of 0.4 μ M TEC with 14 μ l of 400 μ M NTP. Both TEC and NTP solutions were prepared in TB10 buffer. The reaction was allowed to proceed for 0.004–10 s at 25°C, quenched with 86 μ l of 0.5 M HCl and immediately neutralized by adding 171 μ l of neutralizing-loading buffer (94% formamide, 290 mM Tris base, 13 mM Li₄-EDTA, 0.2% Orange G). RNA extension was also followed in 6-MI fluorescence assays by withdrawing 8 μ l aliquots from the fluorometer cuvette into 12 μ l of gel loading buffer (94% formamide, 20 mM Li₄-EDTA and 0.2% Orange G). RNA cleavage was monitored by manual mixing of 50 μ l of 0.2 μ M TEC in TBO buffer with 50 μ l of 16 μ M GreA in TB2 buffer. The aliquots (8 μ l) were withdrawn at the indicated time points and quenched with 12 μ l of the gel loading buffer. The TEC solutions were supplemented with 4 μ M NusG where indicated. RNAs were separated on 16% denaturing polyacrylamide gels and visualized with Odyssey Infrared Imager (Li-Cor Biosciences, Lincoln, NE); band intensities were quantified using ImageJ software (Abramoff *et al.*, 2004).

Time resolved measurements with the 6-MI fluorescent beacon

Measurements were performed in an Applied Photophysics (Leatherhead, UK) SX.18 MV stopped-flow instrument at 25°C. 6-MI fluorophore was excited at 340 nm and emitted light was collected through 400 nm longpass filter. At least three individual traces were averaged for each reported curve. The nucleotide addition, pyrophosphorolysis and RNA cleavage reactions were initiated by mixing 60 μ l of 0.2 μ M TEC with 60 μ l of 400 μ M NTP, 1000 μ M PP_i and 16 (or 4) μ M of GreA, respectively. TEC solutions were supplemented with 4 μ M NusG where indicated. The NusG binding reaction was initiated by mixing 60 μ l of 0.4 μ M TEC with 60 μ l of 0.2–20 μ M NusG. In the nucleotide addition and pyrophosphorolysis experiments reactant solutions were prepared in TB10 buffer, whereas in NusG binding experiments TB1 buffer was used. In the RNA cleavage experiments, TEC and GreA solutions were prepared in TBO and TB2 buffers, respectively.

Time resolved measurements of the dinucleotide release

Measurements were performed in an Applied Photophysics SX.18 MV stopped-flow instrument at 25°C. 2-AP fluorophore was excited at 320 nm and emitted light was collected through 375 nm longpass filter. At least three individual traces were averaged for each reported curve. The RNA cleavage reactions were initiated by mixing 60 μ l of 0.2 μ M TEC with 60 μ l of 16 (or 4) μ M of GreA. TEC and

Table 4. TEC assembly ratios and reaction buffers.

	Concentrations during assembly, μ M				Assembly buffer	Reaction buffer*		
	RNA	tDNA	ntDNA	RNAP		TEC	additive	
Nucleotide addition (gel)	1	1.4	2	1.5	TB10	TB10	TB10	<i>Figure 2B</i>
RNA cleavage (gel)	1	1.4	2	1.5	TB0	TB0	TB2	<i>Figure 3B</i>
Forward translocation (nucleotide addition)	1.4	1	2	1.5	TB10	TB10	TB10	<i>Figure 2B</i>
Backward translocation (pyrophosphorolysis)	1.4	1	2	1.5	TB10	TB10	TB10	<i>Figure 2C</i>
Backtracking (RNA cleavage)	1.4	1	2	1.5	TB0	TB0	TB2	<i>Figure 3B</i>
NusG binding	1.4	1	2	1	TB1	TB1	TB1	<i>Figure 6C</i>
Equilibrium 6-MI assays	1.4	1	2	1.5	TB10	TB10		<i>Figures 2D,6–7</i>
Dinucleotide release (RNA cleavage)	1	1.4	2	1.5	TB0	TB0	TB2	<i>Figure 3</i>
8-MP crosslinking	1.2	1	1	1.5	TB10	TB10		<i>Figures 4,7A</i>
6-TG crosslinking	1	1	2	1.5	TB10	TB10		<i>Figure 5</i>

* In time resolved assays the equal volumes of the TEC and the additive solutions were mixed to initiate the reaction.

DOI: 10.7554/eLife.18096.025

Table 5. The number of repeats for each experiment.

Figure	Data	Number of experiments					
		with independently assembled TECs		including the experiments with the same TEC preparation		with independently assembled TECs not in the figures	
		control	+NusG	control	+NusG	control	+NusG
2B	WT catalysis	4	3	8	6		
	WT translocation	3	2	>12	>8		
2C	WT pyrophosphorolysis	2	2	>8	>8		
2D	WT TGT binding	2	2				
3BC	WT RNA cleavage	2	2	3	4	6	2
	WT 6-MI	2	2	>8	>8		
	WT 2-AP	2	2	>8	>8		
	Δ RL 2-AP	2	2	>8	>8		
	Δ LL 2-AP	2	2	>8	>8		
	Δ GL 2-AP	2	2	>8	>8		
	WT mm1 2-AP	2	2	>8	>8		
	WT mm1-2 2-AP	2	2	>8	>8		
	WT 3'mm 2-AP	2	2	>8	>8		
	3S3	WT 6-MI	2	2	5	5	1
WT 2-AP		1	1	3	3		
WT mm1-2 AA 2-AP		1	1	7	8		
WT mm1-2 GG 2-AP		2	2	>8	>8		
<p>The experiments reported in the figures were performed with the same batch of GreA. The older and the newer experiments cannot be directly combined with the reported experiments due to the variations in the specific activity of the GreA preparations. However, the relative effect of NusG on the reactions involving backtracking can be estimated from all available data. In the WT TEC NusG inhibits reactions that involve backtracking: 2.61 ± 0.28 fold ($n = 5$, $2 \mu\text{M}$ GreA); 2.62 ± 0.22 fold ($n = 6$, $8 \mu\text{M}$ GreA)</p>							
4AB	WT 8-MP	7	3				
	Δ RL 8-MP	3	2				
	Δ LL 8-MP	3	2				
	WT TEC18+TGT 8-MP	2					
5AB	WT 6-TG	7	3				
5S2	Δ RL 6-TG	2	2				
	Δ LL 6-TG	3	2				
	WT TEC18+TGT 6-TG	2					
6AB,C(Left)	WT 6-MI	2	2	>5	>3	>7	>7
	Δ RL 6-MI	2	2	5*	3*		
	Δ LL 6-MI	2	2	5*	>3		
<p>* Except TEC19. SD of the fluorescence measurements with the same TEC preparation = $2.3 \pm 1.3\%$ ($n = 122$). SD of the fluorescence measurements with the independently assembled TECs = 16% ($n = 8$, WT TEC17 measured with different batches of the fluorescent oligonucleotides). Accordingly, the primary data from all fluorescent experiments cannot be directly combined with a figure but some NusG effects can be estimated with the highest accuracy and precision from all available data. NusG effects on the fluorescence intensity of the WT TECs: $(\text{TEC17NusG} - \text{TEC16NusG}) / (\text{TEC17} - \text{TEC16}) = 1.45 \pm 0.09$ ($n = 9$) $(\text{TEC18NusG} - \text{TEC16NusG}) / (\text{TEC18} - \text{TEC16}) = 0.99 \pm 0.02$ ($n = 7$) $(\text{TEC19NusG} - \text{TEC16NusG}) / (\text{TEC19} - \text{TEC16}) = 0.50 \pm 0.35$ ($n = 3$)</p>							
6C (Right)	WT+NusG 6-MI	2		>6			
7A	TEC16-19 6-MI	2		9 (except TEC19)			
	TEC16-19 6-MI	2		4 (except TEC19)			
	TEC16-18 8-MP	2		2			

DOI: 10.7554/eLife.18096.026

GreA solutions were prepared in TB0 and TB2 buffers, respectively. TEC solutions were supplemented with $4 \mu\text{M}$ NusG where indicated.

Equilibrium measurements with the 6-MI fluorescent beacon

Equilibrium levels of fluorescence were determined by continuously recording light emission at 420 nm (excitation at 340 nm) with an LS-55 spectrofluorometer (PerkinElmer, Waltham, MA) in a 16.160-F/Q/10 quartz cuvette (Starna) at 25°C. The assembled TECs were diluted at 100 nM into 200 μ l of TB10 buffer and the NTP substrates (5 μ M) and/or the increasing concentrations of TGT (where indicated) were sequentially added into the cuvette. The reaction was allowed to proceed for up to two minutes between each successive addition to ensure that the fluorescence reached the equilibrium level.

8-methoxypsoralen (8-MP) and 6-thioguanine (6-TG) crosslinking

In 8-MP (mono-adduct absorption maximum 342 nm [Tessman *et al.*, 1985]) crosslinking experiments the reaction mixture contained 1 μ M TEC, 0.92 mM 8-MP, 6.3% DMSO in TB10 buffer. In 6-TG (absorption maximum 340 nm [Karran and Attard, 2008]) crosslinking experiments the reaction mixture contained 1 μ M TEC in TB10 buffer. NTPs (5 μ M), TGT (5 μ M) and NusG (2.5 μ M) were added were indicated. TEC samples (5 μ l) were placed in an 18-well circular tray (\varnothing =26 mm, all wells equidistant from the center) in a closed thermally controlled chamber with the UV LED (P8D1 365 nm, Seoul Viosys, Ansan, Korea) in the top center (height=17 mm). Samples were exposed to UV for 30 min at 25°C, 4 μ l aliquots were quenched with 6 μ l of loading buffer and separated on 14% denaturing PAGE gel. ATTO680 labeled species were visualized with Odyssey Infrared Imager (Li-Cor Biosciences, Lincoln, NE); band intensities were quantified using ImageJ software (Abramoff *et al.*, 2004).

TEC-NusG model

The composite model was generated using the structure of *T. thermophilus* TEC with the NTP analogue (Vassilyev *et al.*, 2007) (PDB ID 2O5J, the lineage specific domain (β '132–456) omitted), NusG NTD from the model of *T. thermophilus* NusG-RNAP complex in Martinez-Rucobo *et al.* (2011), NusG CTD (G187-I248) from the crystal structure of *Aquifex aeolicus* NusG (Steiner *et al.*, 2002) (PDB ID 1M1G) and α CTDs from the crystal structure of *E. coli* holoenzyme (Murakami, 2013) (PDB ID 4YG2). The duplex DNA immediately upstream of the RNA:DNA hybrid was modeled de novo as described in results; the downstream DNA outside of the TEC was extended with the canonical DNA duplex. The positions of NusG CTD and α CTD were chosen arbitrary within the volume permitted by the length of the flexible linkers tethering those domains to the TEC. Parts of the linkers were modeled de novo using ModLoop RRID:SCR_008395 (Fiser and Sali, 2003). NusG CTD and α CTDs are highly conserved in bacteria but are likely irrelevant to the NusG effects in the present study. The above considerations justify the use of heterologous NusG CTD and α CTD in the composite model solely for the illustrative purposes. The model geometry was evaluated using MolProbity RRID:SCR_014226 (Chen *et al.*, 2010). The atomic coordinates of the TEC-NusG complex are provided as Figure 8—source data 1 (NusG-TEC.pdb). To generate Figures 1 and 8A the simplified surfaces (Gaussian resolution 6, B-factor 50) were calculated and rendered in PyMOL Molecular Graphics System, RRID:SCR_000305, (Schrödinger, New York, NY), exported in VRML format, converted to OBJ format using Meshlab and further simplified using sculpting tools of Meshmixer (Autodesk Inc. San Rafael, CA). The resulting meshes were imported into and rendered in Rhinoceros 4.0 RRID:SCR_014339 (Robert McNeel & Associates, Seattle, WA). The superimposition of *Aquifex aeolicus* NusG NTD (PDB ID 1M1G, residues A9-50, A133-185) and *Pyrococcus furiosus* Spt5 NTD (PDB ID 3P8B, residues B4-82) was performed using COLORBYRMSD PyMOL plugin (by S. Shandilya, J. Vertrees, T. Holder).

Data analyses

Time-resolved nucleotide incorporation and the forward translocation data were simultaneously fit to a three-step model using the numerical integration capabilities of KinTek Explorer software (Johnson, 2009) (KinTek Corporation, Austin, TX). The model postulated that the initial TEC16 slowly and reversibly interconverts between inactive and active states and, upon the addition of the NTP substrate, undergoes an irreversible transition to TEC17, followed by irreversible translocation (Supplementary file 1) (Malinen *et al.*, 2014). Pyrophosphorolysis, backtracking, RNA cleavage and dinucleotide release data were fit to the stretched exponential function and the median reaction

times were used in place of half-lives to quantify the reaction progress (**Supplementary file 2**). Equilibrium titration data were fit to the dissociation equilibrium equations that accounted for changes in concentrations of all reactants upon complex formation using Scientist 2.01 software (Micromath, Saint Louis, MO) (**Supplementary file 3**). NusG binding data were fit to a one-step reversible binding model (**Supplementary file 4**). Numerical values of the reaction rate constants and the median reaction times are presented in **Table 1**. The number of repeats for each experiment is indicated in **Table 5**.

Acknowledgements

The authors would like to thank Irina Artsimovitch for critically reading the manuscript and providing expression plasmids; Thadée Grocholski, Henri Malmi, Salli Keinänen and Pavlína Gregorova for assistance with protein purification, cloning and biochemical experiments; Anssi M Malinen for helpful discussions and Jani Sointusalo for making the crosslinking device. Essential equipment was contributed by Walter and Lisi Wahl Foundation.

Additional information

Funding

Funder	Grant reference number	Author
Suomen Akatemia	grant #286205	Matti Turtola Georgiy A Belogurov
Turun Yliopisto	Graduate Student Fellowship	Matti Turtola

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Author contributions

MT, Designed and performed the biochemical experiments, Interpreted the results, Wrote the manuscript; GAB, Supervised the study, Performed structural modeling, Interpreted the results, Wrote the manuscript, Conception and design

Author ORCIDs

Matti Turtola,  <http://orcid.org/0000-0003-1694-1027>

Georgiy A Belogurov,  <http://orcid.org/0000-0002-3070-6843>

Additional files

Supplementary files

- Supplementary file 1. Model employed for fitting of nucleotide addition and translocation data.
DOI: [10.7554/eLife.18096.027](https://doi.org/10.7554/eLife.18096.027)
- Supplementary file 2. Model employed for fitting pyrophosphorolysis, backtracking, RNA cleavage and dinucleotide release data.
DOI: [10.7554/eLife.18096.028](https://doi.org/10.7554/eLife.18096.028)
- Supplementary file 3. Model employed for fitting equilibrium titration with TGT.
DOI: [10.7554/eLife.18096.029](https://doi.org/10.7554/eLife.18096.029)
- Supplementary file 4. Model employed for fitting NusG binding kinetics.
DOI: [10.7554/eLife.18096.030](https://doi.org/10.7554/eLife.18096.030)
- Supplementary file 5. Plasmid sequences.
DOI: [10.7554/eLife.18096.031](https://doi.org/10.7554/eLife.18096.031)

References

- Abramoff MD**, Magalhaes PJ, Ram SJ. 2004. *Image Processing with ImageJ*. Biophotonics International. p 36–42.
- Adelman K**, Lis JT. 2012. Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. *Nature Reviews Genetics* **13**:720–731. doi: [10.1038/nrg3293](https://doi.org/10.1038/nrg3293), PMID: [22986266](https://pubmed.ncbi.nlm.nih.gov/22986266/)
- Andrecka J**, Treutlein B, Arcusa MA, Muschielok A, Lewis R, Cheung AC, Cramer P, Michaelis J. 2009. Nano positioning system reveals the course of upstream and nontemplate DNA within the RNA polymerase II elongation complex. *Nucleic Acids Research* **37**:5803–5809. doi: [10.1093/nar/gkp601](https://doi.org/10.1093/nar/gkp601), PMID: [19620213](https://pubmed.ncbi.nlm.nih.gov/19620213/)
- Anthony LC**, Artsimovitch I, Svetlov V, Landick R, Burgess RR. 2000. Rapid purification of His(6)-tagged *Bacillus subtilis* core RNA polymerase. *Protein Expression and Purification* **19**:350–354. doi: [10.1006/prep.2000.1272](https://doi.org/10.1006/prep.2000.1272), PMID: [10910724](https://pubmed.ncbi.nlm.nih.gov/10910724/)
- Arimbasseri AG**, Maraia RJ. 2015. Mechanism of transcription termination by RNA polymerase III utilizes a non-template strand sequence-specific signal element. *Molecular Cell* **58**:1124–1132. doi: [10.1016/j.molcel.2015.04.002](https://doi.org/10.1016/j.molcel.2015.04.002), PMID: [25959395](https://pubmed.ncbi.nlm.nih.gov/25959395/)
- Artsimovitch I**, Landick R. 2000. Pausing by bacterial RNA polymerase is mediated by mechanistically distinct classes of signals. *PNAS* **97**:7090–7095. doi: [10.1073/pnas.97.13.7090](https://doi.org/10.1073/pnas.97.13.7090), PMID: [10860976](https://pubmed.ncbi.nlm.nih.gov/10860976/)
- Artsimovitch I**, Landick R. 2002. The transcriptional regulator RfaH stimulates RNA chain synthesis after recruitment to elongation complexes by the exposed nontemplate DNA strand. *Cell* **109**:193–203. doi: [10.1016/S0092-8674\(02\)00724-9](https://doi.org/10.1016/S0092-8674(02)00724-9), PMID: [12007406](https://pubmed.ncbi.nlm.nih.gov/12007406/)
- Bae B**, Feklistov A, Lass-Napiorkowska A, Landick R, Darst SA. 2015. Structure of a bacterial RNA polymerase holoenzyme open promoter complex. *eLife* **4**:e08504. doi: [10.7554/eLife.08504](https://doi.org/10.7554/eLife.08504)
- Bar-Nahum G**, Epshtein V, Ruckenstein AE, Rafikov R, Mustaev A, Nudler E. 2005. A ratchet mechanism of transcription elongation and its control. *Cell* **120**:183–193. doi: [10.1016/j.cell.2004.11.045](https://doi.org/10.1016/j.cell.2004.11.045), PMID: [15680325](https://pubmed.ncbi.nlm.nih.gov/15680325/)
- Barnes CO**, Calero M, Malik I, Graham BW, Spahr H, Lin G, Cohen AE, Brown IS, Zhang Q, Pullara F, Trakselis MA, Kaplan CD, Calero G. 2015. Crystal structure of a transcribing RNA polymerase II complex reveals a complete transcription bubble. *Molecular Cell* **59**:258–269. doi: [10.1016/j.molcel.2015.06.034](https://doi.org/10.1016/j.molcel.2015.06.034), PMID: [26186291](https://pubmed.ncbi.nlm.nih.gov/26186291/)
- Belogurov GA**, Artsimovitch I. 2015. Regulation of transcript elongation. *Annual Review of Microbiology* **69**:49–69. doi: [10.1146/annurev-micro-091014-104047](https://doi.org/10.1146/annurev-micro-091014-104047), PMID: [26132790](https://pubmed.ncbi.nlm.nih.gov/26132790/)
- Belogurov GA**, Sevostyanova A, Svetlov V, Artsimovitch I. 2010. Functional regions of the N-terminal domain of the antiterminator RfaH. *Molecular Microbiology* **76**:286–301. doi: [10.1111/j.1365-2958.2010.07056.x](https://doi.org/10.1111/j.1365-2958.2010.07056.x), PMID: [20132437](https://pubmed.ncbi.nlm.nih.gov/20132437/)
- Belogurov GA**, Vassilyeva MN, Svetlov V, Klyuyev S, Grishin NV, Vassilyev DG, Artsimovitch I. 2007. Structural basis for converting a general transcription factor into an operon-specific virulence regulator. *Molecular Cell* **26**:117–129. doi: [10.1016/j.molcel.2007.02.021](https://doi.org/10.1016/j.molcel.2007.02.021), PMID: [17434131](https://pubmed.ncbi.nlm.nih.gov/17434131/)
- Bernecky C**, Herzog F, Baumeister W, Plitzko JM, Cramer P. 2016. Structure of transcribing mammalian RNA polymerase II. *Nature* **529**:551–554. doi: [10.1038/nature16482](https://doi.org/10.1038/nature16482), PMID: [26789250](https://pubmed.ncbi.nlm.nih.gov/26789250/)
- Brueckner F**, Cramer P. 2008. Structural basis of transcription inhibition by alpha-amanitin and implications for RNA polymerase II translocation. *Nature Structural & Molecular Biology* **15**:811–818. doi: [10.1038/nsmb.1458](https://doi.org/10.1038/nsmb.1458), PMID: [18552824](https://pubmed.ncbi.nlm.nih.gov/18552824/)
- Bubunenko M**, Court DL, Al Refaii A, Saxena S, Korepanov A, Friedman DI, Gottesman ME, Alix JH. 2013. Nus transcription elongation factors and RNase III modulate small ribosome subunit biogenesis in *Escherichia coli*. *Molecular Microbiology* **87**:382–393. doi: [10.1111/mmi.12105](https://doi.org/10.1111/mmi.12105), PMID: [23190053](https://pubmed.ncbi.nlm.nih.gov/23190053/)
- Burmam BM**, Rösch P. 2011. The role of *E. coli* Nus-factors in transcription regulation and transcription: translation coupling: From structure to mechanism. *Transcription* **2**:130–134. doi: [10.4161/trns.2.3.15671](https://doi.org/10.4161/trns.2.3.15671), PMID: [21922055](https://pubmed.ncbi.nlm.nih.gov/21922055/)
- Burmam BM**, Schweimer K, Luo X, Wahl MC, Stitt BL, Gottesman ME, Rösch P. 2010. A NusE:NusG complex links transcription and translation. *Science* **328**:501–504. doi: [10.1126/science.1184953](https://doi.org/10.1126/science.1184953), PMID: [20413501](https://pubmed.ncbi.nlm.nih.gov/20413501/)
- Burova E**, Hung SC, Sagitov V, Stitt BL, Gottesman ME. 1995. *Escherichia coli* NusG protein stimulates transcription elongation rates in vivo and in vitro. *Journal of Bacteriology* **177**:1388–1392. PMID: [7868616](https://pubmed.ncbi.nlm.nih.gov/7868616/)
- Cardinale CJ**, Washburn RS, Tadigotla VR, Brown LM, Gottesman ME, Nudler E. 2008. Termination factor Rho and its cofactors NusA and NusG silence foreign DNA in *E. coli*. *Science* **320**:935–938. doi: [10.1126/science.1152763](https://doi.org/10.1126/science.1152763), PMID: [18487194](https://pubmed.ncbi.nlm.nih.gov/18487194/)
- Chen VB**, Arendall WB, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC. 2010. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallographica Section D Biological Crystallography* **66**:12–21. doi: [10.1107/S0907444909042073](https://doi.org/10.1107/S0907444909042073), PMID: [20057044](https://pubmed.ncbi.nlm.nih.gov/20057044/)
- Coban O**, Lamb DC, Zaychikov E, Heumann H, Nienhaus GU. 2006. Conformational heterogeneity in RNA polymerase observed by single-pair FRET microscopy. *Biophysical Journal* **90**:4605–4617. doi: [10.1529/biophysj.105.078840](https://doi.org/10.1529/biophysj.105.078840), PMID: [16581837](https://pubmed.ncbi.nlm.nih.gov/16581837/)
- Crickard JB**, Fu J, Reese JC. 2016. Biochemical analysis of yeast suppressor of Ty 4/5 (Spt4/5) reveals the importance of nucleic acid interactions in the prevention of RNA polymerase II Arrest. *Journal of Biological Chemistry* **291**:9853–9870. doi: [10.1074/jbc.M116.716001](https://doi.org/10.1074/jbc.M116.716001), PMID: [26945063](https://pubmed.ncbi.nlm.nih.gov/26945063/)
- Czyz A**, Mooney RA, 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. doi: [10.1128/mBio.00931-14](https://doi.org/10.1128/mBio.00931-14), PMID: [24713321](https://pubmed.ncbi.nlm.nih.gov/24713321/)
- D'Heygère F**, Rabhi M, Boudvillain M. 2013. Phyletic distribution and conservation of the bacterial transcription termination factor Rho. *Microbiology* **159**:1423–1436. doi: [10.1099/mic.0.067462-0](https://doi.org/10.1099/mic.0.067462-0), PMID: [23704790](https://pubmed.ncbi.nlm.nih.gov/23704790/)

- Deaconescu AM**, Chambers AL, Smith AJ, Nickels BE, Hochschild A, Savery NJ, Darst SA. 2006. Structural basis for bacterial transcription-coupled DNA repair. *Cell* **124**:507–520. doi: [10.1016/j.cell.2005.11.045](https://doi.org/10.1016/j.cell.2005.11.045), PMID: [16469698](https://pubmed.ncbi.nlm.nih.gov/16469698/)
- Epshtein V**, Kamarthapu V, McGary K, Svetlov V, Ueberheide B, Proshkin S, Mironov A, Nudler E. 2014. UvrD facilitates DNA repair by pulling RNA polymerase backwards. *Nature* **505**:372–377. doi: [10.1038/nature12928](https://doi.org/10.1038/nature12928), PMID: [24402227](https://pubmed.ncbi.nlm.nih.gov/24402227/)
- Fiser A**, Sali A. 2003. ModLoop: automated modeling of loops in protein structures. *Bioinformatics* **19**:2500–2501. doi: [10.1093/bioinformatics/btg362](https://doi.org/10.1093/bioinformatics/btg362), PMID: [14668246](https://pubmed.ncbi.nlm.nih.gov/14668246/)
- Hein PP**, Kolb KE, Windgassen T, Bellecourt MJ, Darst SA, Mooney RA, Landick R. 2014. RNA polymerase pausing and nascent-RNA structure formation are linked through clamp-domain movement. *Nature Structural & Molecular Biology* **21**:794–802. doi: [10.1038/nsmb.2867](https://doi.org/10.1038/nsmb.2867), PMID: [25108353](https://pubmed.ncbi.nlm.nih.gov/25108353/)
- Herbert KM**, Zhou J, Mooney RA, Porta AL, Landick R, Block SM. 2010. E. coli NusG inhibits backtracking and accelerates pause-free transcription by promoting forward translocation of RNA polymerase. *Journal of Molecular Biology* **399**:17–30. doi: [10.1016/j.jmb.2010.03.051](https://doi.org/10.1016/j.jmb.2010.03.051), PMID: [20381500](https://pubmed.ncbi.nlm.nih.gov/20381500/)
- Hirtreiter A**, Damsma GE, Cheung AC, Klose D, Grohmann D, Vojnic E, Martin AC, Cramer P, Werner F. 2010. Spt4/5 stimulates transcription elongation through the RNA polymerase clamp coiled-coil motif. *Nucleic Acids Research* **38**:4040–4051. doi: [10.1093/nar/gkq135](https://doi.org/10.1093/nar/gkq135), PMID: [20197319](https://pubmed.ncbi.nlm.nih.gov/20197319/)
- Ingham CJ**, Furneaux PA. 2000. Mutations in the ss subunit of the Bacillus subtilis RNA polymerase that confer both rifampicin resistance and hypersensitivity to NusG. *Microbiology* **146**:3041–3049. doi: [10.1099/00221287-146-12-3041](https://doi.org/10.1099/00221287-146-12-3041), PMID: [11101662](https://pubmed.ncbi.nlm.nih.gov/11101662/)
- Johnson KA**. 2009. Fitting enzyme kinetic data with KinTek Global Kinetic Explorer. *Methods in Enzymology* **467**:601–626. doi: [10.1016/S0076-6879\(09\)67023-3](https://doi.org/10.1016/S0076-6879(09)67023-3), PMID: [19897109](https://pubmed.ncbi.nlm.nih.gov/19897109/)
- Karran P**, Attard N. 2008. Thiopurines in current medical practice: molecular mechanisms and contributions to therapy-related cancer. *Nature Reviews Cancer* **8**:24–36. doi: [10.1038/nrc2292](https://doi.org/10.1038/nrc2292), PMID: [18097462](https://pubmed.ncbi.nlm.nih.gov/18097462/)
- Kireeva ML**, Kashlev M. 2009. Mechanism of sequence-specific pausing of bacterial RNA polymerase. *PNAS* **106**:8900–8905. doi: [10.1073/pnas.0900407106](https://doi.org/10.1073/pnas.0900407106), PMID: [19416863](https://pubmed.ncbi.nlm.nih.gov/19416863/)
- Klein BJ**, Bose D, Baker KJ, Yusoff ZM, Zhang X, Murakami KS. 2011. RNA polymerase and transcription elongation factor Spt4/5 complex structure. *PNAS* **108**:546–550. doi: [10.1073/pnas.1013828108](https://doi.org/10.1073/pnas.1013828108), PMID: [21187417](https://pubmed.ncbi.nlm.nih.gov/21187417/)
- Kolb KE**, Hein PP, Landick R. 2014. Antisense oligonucleotide-stimulated transcriptional pausing reveals RNA exit channel specificity of RNA polymerase and mechanistic contributions of NusA and RfaH. *Journal of Biological Chemistry* **289**:1151–1163. doi: [10.1074/jbc.M113.521393](https://doi.org/10.1074/jbc.M113.521393), PMID: [24275665](https://pubmed.ncbi.nlm.nih.gov/24275665/)
- Komissarova N**, Kashlev M. 1997. Transcriptional arrest: Escherichia coli RNA polymerase translocates backward, leaving the 3' end of the RNA intact and extruded. *PNAS* **94**:1755–1760. doi: [10.1073/pnas.94.5.1755](https://doi.org/10.1073/pnas.94.5.1755), PMID: [9050851](https://pubmed.ncbi.nlm.nih.gov/9050851/)
- Komissarova N**, Kireeva ML, Becker J, Sidorenkov I, Kashlev M. 2003. Engineering of elongation complexes of bacterial and yeast RNA polymerases. *Methods in Enzymology* **371**:233–251. doi: [10.1016/S0076-6879\(03\)71017-9](https://doi.org/10.1016/S0076-6879(03)71017-9), PMID: [14712704](https://pubmed.ncbi.nlm.nih.gov/14712704/)
- Kramer NJ**, Carlomagno Y, Zhang YJ, Almeida S, Cook CN, Gendron TF, Prudencio M, Van Blitterswijk M, Belzil V, Couthouis J, Paul JW, Goodman LD, Daugherty L, Chew J, Garrett A, Pregent L, Jansen-West K, Tabassian LJ, Rademakers R, Boylan K, et al. 2016. Spt4 selectively regulates the expression of C9orf72 sense and antisense mutant transcripts. *Science* **353**:708–712. doi: [10.1126/science.aaf7791](https://doi.org/10.1126/science.aaf7791), PMID: [27516603](https://pubmed.ncbi.nlm.nih.gov/27516603/)
- Leela JK**, Syeda AH, Anupama K, Gowrishankar J, Krishna Leela J. 2013. Rho-dependent transcription termination is essential to prevent excessive genome-wide R-loops in Escherichia coli. *PNAS* **110**:258–263. doi: [10.1073/pnas.1213123110](https://doi.org/10.1073/pnas.1213123110), PMID: [23251031](https://pubmed.ncbi.nlm.nih.gov/23251031/)
- Liu B**, Zuo Y, Steitz TA. 2016. Structures of E. coli σ S-transcription initiation complexes provide new insights into polymerase mechanism. *PNAS* **113**:4051–4056. doi: [10.1073/pnas.1520555113](https://doi.org/10.1073/pnas.1520555113), PMID: [27035955](https://pubmed.ncbi.nlm.nih.gov/27035955/)
- Malinen AM**, Nandymazumdar M, Turtola M, Malmi H, Grocholski T, Artsimovitch I, Belogurov GA. 2014. CBR antimicrobials alter coupling between the bridge helix and the β subunit in RNA polymerase. *Nature Communications* **5**:3408. doi: [10.1038/ncomms4408](https://doi.org/10.1038/ncomms4408), PMID: [24598909](https://pubmed.ncbi.nlm.nih.gov/24598909/)
- Malinen AM**, Turtola M, Belogurov GA. 2015. Monitoring translocation of multisubunit RNA polymerase along the DNA with fluorescent base analogues. *Methods in Molecular Biology* **1276**:31–51. doi: [10.1007/978-1-4939-2392-2_3](https://doi.org/10.1007/978-1-4939-2392-2_3), PMID: [25665557](https://pubmed.ncbi.nlm.nih.gov/25665557/)
- Malinen AM**, Turtola M, Parthiban M, Vainonen L, Johnson MS, Belogurov GA. 2012. Active site opening and closure control translocation of multisubunit RNA polymerase. *Nucleic Acids Research* **40**:7442–7451. doi: [10.1093/nar/gks383](https://doi.org/10.1093/nar/gks383), PMID: [22570421](https://pubmed.ncbi.nlm.nih.gov/22570421/)
- Martinez-Rucobo FW**, Sainsbury S, Cheung AC, Cramer P. 2011. Architecture of the RNA polymerase-Spt4/5 complex and basis of universal transcription processivity. *The EMBO Journal* **30**:1302–1310. doi: [10.1038/emboj.2011.64](https://doi.org/10.1038/emboj.2011.64), PMID: [21386817](https://pubmed.ncbi.nlm.nih.gov/21386817/)
- McGary K**, Nudler E. 2013. RNA polymerase and the ribosome: the close relationship. *Current Opinion in Microbiology* **16**:112–117. doi: [10.1016/j.mib.2013.01.010](https://doi.org/10.1016/j.mib.2013.01.010), PMID: [23433801](https://pubmed.ncbi.nlm.nih.gov/23433801/)
- Mooney RA**, Davis SE, Peters JM, Rowland JL, Ansari AZ, Landick R. 2009a. Regulator trafficking on bacterial transcription units in vivo. *Molecular Cell* **33**:97–108. doi: [10.1016/j.molcel.2008.12.021](https://doi.org/10.1016/j.molcel.2008.12.021)
- Mooney RA**, Schweimer K, Rösch P, Gottesman M, Landick R. 2009b. 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. doi: [10.1016/j.jmb.2009.05.078](https://doi.org/10.1016/j.jmb.2009.05.078)

- Murakami KS.** 2013. X-ray crystal structure of Escherichia coli RNA polymerase $\sigma 70$ holoenzyme. *Journal of Biological Chemistry* **288**:9126–9160. doi: [10.1074/jbc.M112.430900](https://doi.org/10.1074/jbc.M112.430900), PMID: [23389035](https://pubmed.ncbi.nlm.nih.gov/23389035/)
- Nudler E, Mustaev A, Lukhtanov E, Goldfarb A.** 1997. The RNA-DNA hybrid maintains the register of transcription by preventing backtracking of RNA polymerase. *Cell* **89**:33–41. doi: [10.1016/S0092-8674\(00\)80180-4](https://doi.org/10.1016/S0092-8674(00)80180-4), PMID: [9094712](https://pubmed.ncbi.nlm.nih.gov/9094712/)
- Nudler E.** 2012. RNA polymerase backtracking in gene regulation and genome instability. *Cell* **149**:1438–1445. doi: [10.1016/j.cell.2012.06.003](https://doi.org/10.1016/j.cell.2012.06.003), PMID: [22726433](https://pubmed.ncbi.nlm.nih.gov/22726433/)
- Opalka N, Brown J, Lane WJ, Twist KA, Landick R, Asturias FJ, Darst SA.** 2010. Complete structural model of Escherichia coli RNA polymerase from a hybrid approach. *PLoS Biology* **8**:e1000483. doi: [10.1371/journal.pbio.1000483](https://doi.org/10.1371/journal.pbio.1000483), PMID: [20856905](https://pubmed.ncbi.nlm.nih.gov/20856905/)
- 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. doi: [10.1021/bi992658z](https://doi.org/10.1021/bi992658z), PMID: [820031](https://pubmed.ncbi.nlm.nih.gov/820031/)
- Perdue SA, Roberts JW.** 2011. $\Sigma(70)$ -dependent transcription pausing in Escherichia coli. *Journal of Molecular Biology* **412**:782–792. doi: [10.1016/j.jmb.2011.02.011](https://doi.org/10.1016/j.jmb.2011.02.011), PMID: [21316374](https://pubmed.ncbi.nlm.nih.gov/21316374/)
- Perederina AA, Vassilyeva MN, Berezin IA, Svetlov V, Artsimovitch I, Vassilyev DG.** 2006. Cloning, expression, purification, crystallization and initial crystallographic analysis of transcription elongation factors GreB from Escherichia coli and Gfh1 from Thermus thermophilus. *Acta Crystallographica Section F Structural Biology and Crystallization Communications* **62**:44–46. doi: [10.1107/S1744309105040297](https://doi.org/10.1107/S1744309105040297), PMID: [16511259](https://pubmed.ncbi.nlm.nih.gov/16511259/)
- Peters JM, Mooney RA, Grass JA, Jessen ED, Tran F, Landick R.** 2012. Rho and NusG suppress pervasive antisense transcription in Escherichia coli. *Genes & Development* **26**:2621–2633. doi: [10.1101/gad.196741.112](https://doi.org/10.1101/gad.196741.112), PMID: [23207917](https://pubmed.ncbi.nlm.nih.gov/23207917/)
- Peters JM, Mooney RA, Kuan PF, Rowland JL, Keles S, Landick R.** 2009. Rho directs widespread termination of intragenic and stable RNA transcription. *PNAS* **106**:15406–15411. doi: [10.1073/pnas.0903846106](https://doi.org/10.1073/pnas.0903846106), PMID: [19706412](https://pubmed.ncbi.nlm.nih.gov/19706412/)
- Proshkin S, Rahmouni AR, Mironov A, Nudler E.** 2010. Cooperation between translating ribosomes and RNA polymerase in transcription elongation. *Science* **328**:504–508. doi: [10.1126/science.1184939](https://doi.org/10.1126/science.1184939), PMID: [20413502](https://pubmed.ncbi.nlm.nih.gov/20413502/)
- Santangelo TJ, Artsimovitch I.** 2011. Termination and antitermination: RNA polymerase runs a stop sign. *Nature Reviews Microbiology* **9**:319–329. doi: [10.1038/nrmicro2560](https://doi.org/10.1038/nrmicro2560), PMID: [21478900](https://pubmed.ncbi.nlm.nih.gov/21478900/)
- Schulz S, Gietl A, Smollett K, Tinnfeld P, Werner F, Grohmann D.** 2016. TFE and Spt4/5 open and close the RNA polymerase clamp during the transcription cycle. *PNAS* **113**:E1816–E1825. doi: [10.1073/pnas.1515817113](https://doi.org/10.1073/pnas.1515817113), PMID: [26979960](https://pubmed.ncbi.nlm.nih.gov/26979960/)
- Sekine S, Murayama Y, Svetlov V, Nudler E, Yokoyama S.** 2015. The ratcheted and ratchetable structural states of RNA polymerase underlie multiple transcriptional functions. *Molecular Cell* **57**:408–421. doi: [10.1016/j.molcel.2014.12.014](https://doi.org/10.1016/j.molcel.2014.12.014), PMID: [25601758](https://pubmed.ncbi.nlm.nih.gov/25601758/)
- Sevostyanova A, Artsimovitch I.** 2010. Functional analysis of Thermus thermophilus transcription factor NusG. *Nucleic Acids Research* **38**:7432–7445. doi: [10.1093/nar/gkq623](https://doi.org/10.1093/nar/gkq623), PMID: [20639538](https://pubmed.ncbi.nlm.nih.gov/20639538/)
- Sevostyanova A, Belogurov GA, Mooney RA, Landick R, Artsimovitch I.** 2011. The β subunit gate loop is required for RNA polymerase modification by RfaH and NusG. *Molecular Cell* **43**:253–262. doi: [10.1016/j.molcel.2011.05.026](https://doi.org/10.1016/j.molcel.2011.05.026), PMID: [21777814](https://pubmed.ncbi.nlm.nih.gov/21777814/)
- Shankar S, Hatoum A, Roberts JW.** 2007. A transcription antiterminator constructs a NusA-dependent shield to the emerging transcript. *Molecular Cell* **27**:914–927. doi: [10.1016/j.molcel.2007.07.025](https://doi.org/10.1016/j.molcel.2007.07.025), PMID: [17889665](https://pubmed.ncbi.nlm.nih.gov/17889665/)
- Singh N, Bubunenko M, Smith C, Abbott DM, Stringer AM, Shi R, Court DL, Wade JT.** 2016. SuhB associates with Nus factors to facilitate 30S ribosome biogenesis in Escherichia coli. *mBio* **7**:e00114–00116. doi: [10.1128/mBio.00114-16](https://doi.org/10.1128/mBio.00114-16), PMID: [26980831](https://pubmed.ncbi.nlm.nih.gov/26980831/)
- Steiner T, Kaiser JT, Marinkovič S, Huber R, Wahl MC.** 2002. Crystal structures of transcription factor NusG in light of its nucleic acid- and protein-binding activities. *The EMBO Journal* **21**:4641–4694. doi: [10.1093/emboj/cdf455](https://doi.org/10.1093/emboj/cdf455), PMID: [12198166](https://pubmed.ncbi.nlm.nih.gov/12198166/)
- Svetlov V, Artsimovitch I.** 2015. Purification of bacterial RNA polymerase: tools and protocols. *Methods in Molecular Biology* **1276**:13–29. doi: [10.1007/978-1-4939-2392-2_2](https://doi.org/10.1007/978-1-4939-2392-2_2), PMID: [25665556](https://pubmed.ncbi.nlm.nih.gov/25665556/)
- Tessman JW, Isaacs ST, Hearst JE.** 1985. Photochemistry of the furan-side 8-methoxypsoralen-thymidine monoadduct inside the DNA helix. Conversion to diadduct and to pyrone-side monoadduct. *Biochemistry* **24**:1669–1676. doi: [10.1021/bi00328a015](https://doi.org/10.1021/bi00328a015), PMID: [4005221](https://pubmed.ncbi.nlm.nih.gov/4005221/)
- Vassilyev DG, Vassilyeva MN, Zhang J, Palangat M, Artsimovitch I, Landick R.** 2007. Structural basis for substrate loading in bacterial RNA polymerase. *Nature* **448**:163–168. doi: [10.1038/nature05931](https://doi.org/10.1038/nature05931), PMID: [17581591](https://pubmed.ncbi.nlm.nih.gov/17581591/)
- Vvedenskaya IO, Vahedian-Movahed H, Bird JG, Knoblauch JG, Goldman SR, Zhang Y, Ebricht RH, Nickels BE.** 2014. Interactions between RNA polymerase and the "core recognition element" counteract pausing. *Science* **344**:1285–1289. doi: [10.1126/science.1253458](https://doi.org/10.1126/science.1253458), PMID: [24926020](https://pubmed.ncbi.nlm.nih.gov/24926020/)
- Wada T, Takagi T, Yamaguchi Y, Ferdous A, Imai T, Hirose S, Sugimoto S, Yano K, Hartzog GA, Winston F, Buratowski S, Handa H.** 1998. DSIF, a novel transcription elongation factor that regulates RNA polymerase II processivity, is composed of human Spt4 and Spt5 homologs. *Genes & Development* **12**:343–356. doi: [10.1101/gad.12.3.343](https://doi.org/10.1101/gad.12.3.343), PMID: [9450929](https://pubmed.ncbi.nlm.nih.gov/9450929/)
- Wang D, Bushnell DA, Huang X, Westover KD, Levitt M, Kornberg RD.** 2009. Structural basis of transcription: backtracked RNA polymerase II at 3.4 angstrom resolution. *Science* **324**:1203–1206. doi: [10.1126/science.1168729](https://doi.org/10.1126/science.1168729), PMID: [19478184](https://pubmed.ncbi.nlm.nih.gov/19478184/)

- Weixlbaumer A**, Leon K, Landick R, Darst SA. 2013. Structural basis of transcriptional pausing in bacteria. *Cell* **152**:431–441. doi: [10.1016/j.cell.2012.12.020](https://doi.org/10.1016/j.cell.2012.12.020), PMID: [23374340](https://pubmed.ncbi.nlm.nih.gov/23374340/)
- 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. doi: [10.1016/j.jmb.2012.01.031](https://doi.org/10.1016/j.jmb.2012.01.031), PMID: [22306403](https://pubmed.ncbi.nlm.nih.gov/22306403/)
- Yakhnin AV**, Murakami KS, Babitzke P. 2016. NusG Is a sequence-specific RNA polymerase pause factor that binds to the non-template DNA within the paused transcription bubble. *Journal of Biological Chemistry* **291**:5299–5308. doi: [10.1074/jbc.M115.704189](https://doi.org/10.1074/jbc.M115.704189), PMID: [26742846](https://pubmed.ncbi.nlm.nih.gov/26742846/)
- Yakhnin AV**, Yakhnin H, Babitzke P. 2008. Function of the *Bacillus subtilis* transcription elongation factor NusG in hairpin-dependent RNA polymerase pausing in the *trp* leader *PNAS*. **105**:16131–16136. doi: [10.1073/pnas.0808842105](https://doi.org/10.1073/pnas.0808842105), PMID: [18852477](https://pubmed.ncbi.nlm.nih.gov/18852477/)
- Zellars M**, Squires CL. 1999. Antiterminator-dependent modulation of transcription elongation rates by NusB and NusG. *Molecular Microbiology* **32**:1296–1304. doi: [10.1046/j.1365-2958.1999.01442.x](https://doi.org/10.1046/j.1365-2958.1999.01442.x), PMID: [10383769](https://pubmed.ncbi.nlm.nih.gov/10383769/)
- Zhang J**, Palangat M, Landick R. 2010. Role of the RNA polymerase trigger loop in catalysis and pausing. *Nature Structural & Molecular Biology* **17**:99–104. doi: [10.1038/nsmb.1732](https://doi.org/10.1038/nsmb.1732), PMID: [19966797](https://pubmed.ncbi.nlm.nih.gov/19966797/)
- Zhang Y**, Feng Y, Chatterjee S, Tuske S, Ho MX, Arnold E, Ebright RH. 2012. Structural basis of transcription initiation. *Science* **338**:1076–1080. doi: [10.1126/science.1227786](https://doi.org/10.1126/science.1227786), PMID: [23086998](https://pubmed.ncbi.nlm.nih.gov/23086998/)
- Zuo Y**, Steitz TA. 2015. Crystal structures of the *E. coli* transcription initiation complexes with a complete bubble. *Molecular Cell* **58**:534–540. doi: [10.1016/j.molcel.2015.03.010](https://doi.org/10.1016/j.molcel.2015.03.010), PMID: [25866247](https://pubmed.ncbi.nlm.nih.gov/25866247/)