7432–7445 *Nucleic Acids Research, 2010, Vol. 38, No. 21* *Published online 17 July 2010*
doi:10.1093/nar/gkq623

Functional analysis of *Thermus thermophilus* transcription factor NusG

Anastasiya Sevostyanova and Irina Artsimovitch*

Department of Microbiology and The Center for RNA Biology, The Ohio State University, Columbus, OH 43210, USA

Received March 22, 2010; Revised June 19, 2010; Accepted June 28, 2010

ABSTRACT

Transcription elongation factors from the NusG family are ubiquitous from bacteria to humans and play diverse roles in the regulation of gene expression. These proteins consist of at least two domains. The N-terminal domains directly bind to the largest, β′ in bacteria, subunit of RNA polymerase (RNAP), whereas the C-terminal domains interact with other cellular components and serve as platforms for the assembly of large nucleoprotein complexes. *Escherichia coli* NusG and its paralog RfaH modify RNAP into a fast, pause-resistant state but the detailed molecular mechanism of this modification remains unclear since no high-resolution structural data are available for the *E. coli* system. We wanted to investigate whether *Thermus thermophilus* (Tth) NusG can be used as a model for structural studies of this family of regulators. Here, we show that Tth NusG slows down rather than facilitates transcript elongation by its cognate RNAP. On the other hand, similarly to the *E. coli* regulators, Tth NusG apparently binds near the upstream end of the transcription bubble, competes with σ^A, and favors forward translocation by RNAP. Our data suggest that the mechanism of NusG recruitment to RNAP is universally conserved even though the regulatory outcomes among its homologs may appear distinct.

INTRODUCTION

The transcription elongation factor NusG has been identified in *Escherichia coli* on the basis of its requirement for phage λ N-dependent gene expression and thus named N utilization substance G (1). Subsequent studies demonstrated that *E. coli* (Eco) NusG affects Rho-dependent termination (2), transcriptional arrest by HK022 Nun protein (3), RNA chain elongation (4) and translation (5), and is also a key component of the rRNA antitermination complex (6). NusG is essential in wild-type *E. coli* (2) and is associated with RNAP transcribing most of the *E. coli* MG1655 genes (7). However, recent studies (8) demonstrate that Eco NusG becomes dispensable when the *rac* prophage *kil* gene is deleted and suggest that the essential role of NusG in *E. coli* is to enhance Rho-dependent termination within the horizontally transferred operons, thereby limiting their expression. In support of this hypothesis, Eco NusG directly interacts with Rho (9) to increase Rho-dependent termination at suboptimal sites (10,11).

Given the variety of functions that have been assigned to NusG, it is difficult to infer which of these functions is (are) the most important. The majority of functional data were collected with Eco NusG (1,4,7,9,11–13); *Bacillus subtilis* (Bsu) (14) and *Thermotoga maritima* (15) NusGs have also been partially characterized. Despite their high sequence and structural conservation, NusG proteins from different bacteria appear to have somewhat different functions and interaction partners. In contrast to *E. coli*, NusG is dispensable in *B. subtilis* (16) and *Staphylococcus aureus* (17) and does not bind to Rho in *T. thermophilus* (18). Many species also encode specialized NusG paralogs. In *E. coli*, RfaH activates expression of LPS and capsule biosynthesis operons, hemolysin, and fertility genes by post-initiation mechanism that has been compared to antitermination (19). In *Bacteroides fragilis,* eight RfaH-like operon-specific UpxY antiterminators regulate expression of capsular polysaccharides (20). The diverse range of binding partners/activities suggests that, after diverging from a common ancestor, different NusG-like proteins became adapted to playing specialized regulatory roles.

Functional analysis of this protein family has been greatly accelerated by the availability of several structures (9,18,21–24) solved by X-ray crystallography and NMR. The N-terminal domains of the NusG-like proteins are similar, whereas their C-terminal domain structures are quite divergent (Figure 1). The differences range from

*To whom correspondence should be addressed. Tel: +1 614 292 6777; Fax: +1 614 292 8120; Email: artsimovitch.1@osu.edu

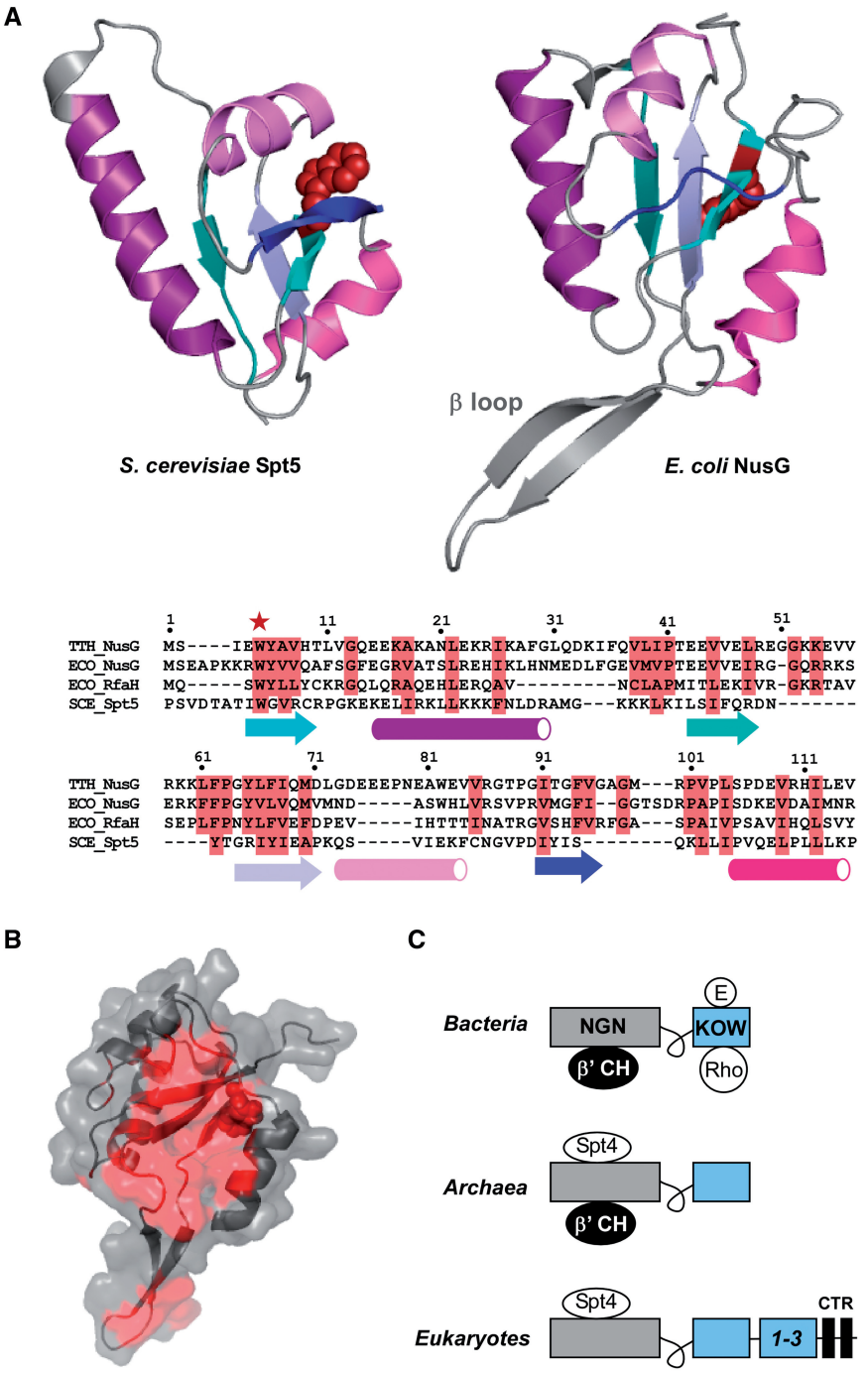


Figure 1. Structural conservation in the NusG family. (**A**) Structural models of the N-terminal domains of *E. coli* NusG (21) and yeast Spt5 (58) proteins. Secondary structure elements are colored and indicated on the sequence alignment shown at the bottom; the side chain of a highly conserved Trp residue is shown as red spheres (and as a red star in the alignment). The conserved residues are indicated in red on the alignment based on multiple sequences; only Tth NusG, Eco NusG, Eco RfaH and yeast Spt5 are shown, the numbering corresponds to Tth NusG. Out of 35 residues conserved among bacterial homologs, the vast majority are hydrophobic (27), six are structural (Pro and Gly). One of two positively charged conserved residues is located in the β loop; deletion of this loop (but not multiple substitutions that remove the charge) eliminates Eco NusG effects on elongation and Rho-dependent termination (39). (**B**) The proposed RNAP-binding surface (23) of Eco NusG lined with conserved hydrophobic residues (including Trp9, spheres) is facing the viewer. (**C**) Domain architecture of NusG-like proteins in the three domains of life; the experimentally confirmed binding partners are shown. NGN, $NusG^{N}$; KOW, Kyprides–Ouzounis–Woese domain; see Steiner *et al.* (21) and references therein; CTR, C-terminal repeats. This figure was made using PyMol (DeLano Scientific, Palo Alto, CA).

the presence of additional domains (Figure 1C) to a complete domain refolding of the C-terminal domain of RfaH ($RfaH^{C}$) relative to that of NusG ($NusG^{C}$). The SH3 β-barrel in NusG is transformed into an α-helical hairpin in RfaH (25), yet the $RfaH^{C}$ sequence can be easily threaded into a NusG-like structure.

Comparison of RfaH and Eco NusG (9,25) shed light on the similarities and differences between their molecular

mechanisms. Like NusG, RfaH increases the apparent rate of RNA chain elongation (19) and reduces pausing and termination at sites where RNAP is prone to backtracking *in vitro*; these effects are likely due to the stabilization of a forward-translocated state (26,27) of the transcription elongation complex (TEC). Unlike NusG, RfaH reduces pausing at hairpin-dependent sites (23), does not bind to Rho (IA, data not shown), and only acts on operons that encode a 12-nt long *ops* element (19). During recruitment to the TEC, RfaH specifically recognizes the *ops* bases in the non-template (NT) DNA strand transiently exposed on the surface of the moving enzyme. This interaction triggers domain dissociation that unmasks the RNAP-binding site located at the domain interface in free RfaH (23). In contrast, the RNAP-binding surface is always exposed in NusG (9), which can be recruited to RNAP at any site on a template. Following their recruitment, NusG (7) and RfaH (25) remain associated with RNAP throughout elongation *in vivo*.

The RfaH and NusG N-terminal domains ($RfaH^N$ and $NusG^N$) are sufficient for their anti-pausing effects and likely bind to the β′ clamp helices domain (β′ CH) of RNAP, as does *Methanococcus jannaschii* (Mja) NusG (28). In fact, even though the isolated $RfaH^N$ recognizes the *ops* sequence during elongation, it no longer requires *ops* for recruitment to the TEC. Thus, the *ops* element serves to restrict the RfaH action to a subset of *E. coli* operons. The C-terminal domains play protein-specific roles: $RfaH^C$ indirectly confers the requirement for the *ops* element (23) and may bind to ribosome (25), $NusG^C$ interacts with Rho (9) and NusE (29) and $Spt5^C$ likely serves as an assembly platform for proteins that promote transcription elongation and histone modification (30).

We are particularly interested in the mechanism of the antipausing modification of RNAP by the NusG-like proteins. These proteins bind ~75 Å away from the RNAP active site (9,23,28) and may act directly, by binding to and stabilizing the upstream DNA fork junction, or allosterically, by altering conformational dynamics of the trigger loop and the bridge helix, the two key catalytic elements in the β′-subunit (31). These mechanisms are not exclusive, and high-resolution structural data will be required to elucidate the fine details of NusG/TEC interactions. Sequence and structural conservation of NusG and its target on RNAP, together with a recent analysis of archaeal NusG (28), suggest that bacterial model systems recapitulate all the aspects of antipausing modification (as opposed to a complex network of regulatory interactions of Spt5 in eukaryotes; see 'Discussion' section). The high-resolution structures of the TEC and the full-length NusG are available only for *T. thermophilus*, which lacks the better-studied RfaH. We wanted to ascertain that Tth NusG can be used as a model for the structure/function analysis. Here, we report that, similarly to RfaH, Tth NusG binds to, and stabilizes the forward translocated state of, the TEC and competes with Tth σ^A during elongation. Thus, the architecture of the NusG-bound TEC appears to be conserved, justifying the choice of Tth NusG as a model for this family of regulators.

MATERIALS AND METHODS

Plasmids and reagents

pTYB12 was obtained from NEB (Ipswich, MA, USA). All general reagents were obtained from Sigma Aldrich (St Louis, MO, USA) and Fisher (Pittsburgh, PA, USA); NTPs, [γ^{32}P]-ATP and [α^{32}P]-GTP, from GE Healthcare (Piscataway, NJ, USA) and Perkin Elmer (Boston, MA, USA); PCR reagents, restriction and modification enzymes, from NEB, Roche (Indianapolis, IN, USA) and Epicentre (Madison, WI, USA). Chitin beads and Ni-sepharose were from NEB and GE Healthcare, respectively. Oligonucleotides were obtained from Integrated DNA Technologies (Coralville, IA, USA) and Sigma Aldrich. DNA purification kits were from Qiagen (Valencia, CA, USA) and Promega (Madison, WI, USA).

Protein expression and purification

XJb (λDE3) strain transformed with pVS58 (Tth NusG ORF cloned between NdeI and NotI sites of pTYB12) was inoculated into LB (Miller) + 0.1 mg/l carbenicillin + auto-induction reagents as described by Studier (32). The culture was grown with agitation at 32°C till stationary phase (~20 h; OD_{600}~5). Arabinose was added to 0.06% after 12–16 h to induce the expression of endolysin. Cells were collected by centrifugation and frozen at −80°C. Pellet was resuspended in IMPACT-CN500 buffer (50 mM Tris–HCl, pH 8.8, 500 mM NaCl, 1 mM EDTA) + 1× Complete® EDTA-free Protease Inhibitors Cocktail (Roche) + 0.1% Tween-20. Cells were lysed by ultrasonication, followed by centrifugation (2× 30 min at 29 500*g*, 4°C), and the cleared lysate was loaded on chitin beads equilibrated with IMPACT-CN500. The column was washed with 10 volumes of IMPACT-CN500, three volumes of IMPACT-CN500 + 50 mM DTT, and incubated at 22°C for 18 h. The cleaved-off protein was eluted with IMPACT-CN500, polished by gel-filtration over a Sephacryl S-200 HR column (GE Healthcare), dialysed against storage buffer (50% glycerol, 100 mM NaCl, 10 mM Tris–HCl pH 7.9, 0.1 mM EDTA, 0.1 mM DTT) and stored at −20°C.

pIA885 containing residues 1–117 of Tth NusG ($NusG^N$) fused to a His_6 tag and a TEV recognition site under the T7 promoter was transformed into XJb (λDE3) strain. An overnight culture was diluted 1/100 into fresh LB (Miller) medium and grown at 37°C. IPTG was added to 1mM at OD_{600} ~0.4, cells were grown for 3.5 h at 30°C, and collected by centrifugation. The pellet was resuspended in Lysis buffer (500 mM NaCl, 50 mM Tris–HCl pH 6.9, 5% glycerol, 0.1 mM EDTA, 1 mM βME, with Complete® cocktail) and disrupted by ultrasonication. The extract was cleared by centrifugation and subjected to heat shock for 20 min at 70°C. Precipitate was removed by centrifugation, supernatant was filtered and loaded onto a Ni–sepharose (GE Healthcare) gravity column preequilibrated with Lysis buffer. Column was washed with 10 volumes of Lysis buffer, 10 volumes of HepA

buffer (50 mM Tris–HCl pH 6.9, 5% glycerol, 1 mM βME) and 10 volumes of HepA+20 mM imidazole. Protein was eluted with HepA+100 mM imidazole and loaded onto a HiTrap Heparin HP column (GE Healthcare). Bound proteins were eluted by NaCl gradient, $NusG^N$ eluted as a single peak at 12 mSi (~220 mM NaCl). $NusG^N$ fractions were concentrated on Amicon filtration device MWCO 5 kDa (Millipore, Billerica, MA, USA) and NaCl concentration was adjusted to 500 mM. His_6 tagged TEV protease (100 μg) was incubated with the protein sample (~8 mg) at 4°C for 48 h. The cleaved-off His_6 tag, the uncut His_6-$NusG^N$, and (His-tagged) TEV were removed by absorption to Ni–sepharose. $NusG^N$ was dialyzed into storage buffer (as described earlier) and stored at −20°C.

Transcript elongation assays

Linear DNA template generated by PCR amplification (30 nM), holo RNAP (40 nM), ApU (100 μM), and starting NTP subsets (1 μM GTP, 5 μM ATP and UTP, 10 μCi [α^{32}P]-GTP, 3000 Ci/mmol) were mixed in 100 μl of TGA10 (20 mM Tris–acetate, 20 mM Na–acetate, 10 mM Mg–acetate, 5% glycerol, 1 mM DTT, 0.1 mM EDTA, pH 7.9). Reactions were incubated for 10 min at 37°C or 55°C for Eco and Tth RNAP, respectively; thus halted TECs were stored on ice. Transcription was restarted by addition of nucleotides (10 μM GTP, 150 μM ATP, CTP and UTP) and heparin to 10 μg/ml at either 37°C or 55°C. Samples were removed at desired time points and quenched by addition of an equal volume of STOP buffer (10 M urea, 60 mM EDTA, 45 mM Tris–borate; pH 8.3).

Pause-free elongation assays

Halted radiolabeled A29 TECs (40 nM) formed on a linear template PCR amplified from pIA146 DNA template containing the pause-less *rpoB* gene in 30 μl of TGA2 buffer with ATP and CTP at 2.5 μM, GTP at 1 μM, ApU at 150 μM and 20 μCi of [α-^{32}P]GTP (3000 Ci/mmol; Perkin Elmer) for 10 min at 55°C, diluted 8-fold and split into six aliquots. Samples were incubated with NusG (200 nM) or storage buffer for 1 min, and transcription was restarted by addition of NTPs (150 μM ATP, CTP, UTP and 10 μM GTP). Aliquots were withdrawn at indicated times, quenched and analyzed on 5% denaturing gels.

Sigma competition assay

Sigma competition assay was performed on a linear pAS33 DNA template (50 nM) containing λ P_R promoter followed by a C-less region, σ-dependent pause (*σP*) and the *hisT* terminator. Halted TECs were prepared in 50 μl of TGA2 buffer (20 mM Tris–acetate, 20 mM Na–acetate, 2 mM Mg–acetate, 5% glycerol, 1 mM DTT, 0.1 mM EDTA, pH 7.9) with holo Tth RNAP (50 nM), ApU (100 μM) and starting NTPs (1 μM GTP, 5 μM ATP and UTP, 10 μCi [α^{32}P]GTP, 3000 Ci/mmol) at 55°C for 10 min. Tth NusG and/or σ^A was added to 0.5 μM followed by a 1-min incubation at 55°C. Transcription was restarted by the addition of all four NTPs to 40 μM and rifapentin to 25 μg/ml.

$KMnO_4$ footprinting

Linear 153-bp DNA fragment containing λ P_R promoter was made by PCR amplification using pIA226 as a template with primers 17 (5′-CGTTAAATCTATCACC GCAAGGG) and 138 (5′-ATCGCCTGAAAGACTAG TCAGG). The NT DNA strand primer (#17) was end-labeled with [^{32}P]-γATP with PNK (Epicentre) and purified using G-50 spin columns (GE Healthcare). PCR products were gel-purified using a Wizard® SV kit (Promega). Sequencing reactions were performed using the same labeled primer with SequiTherm kit (Epicentre). Open complexes were formed with holo Tth RNAP (200 nM) preincubated with the labeled promoter fragment (100 nM) and ApU (100 μM) for 10 min at 55°C in GBB buffer (20 mM Tris–HCl, 20 mM NaCl, 14 mM $MgCl_2$, 5% glycerol and 0.1 mM EDTA; pH 7.9) in the presence of 0.03% DMSO. To form halted complexes, the reaction was supplemented with 1 μM GTP, 5 μM ATP and 5 μM UTP. $NusG^N$ was added to 2 μM where indicated. Samples were shifted to room temperature and treated with 10 mM $KMnO_4$ for 60 s. After addition of an equal volume of quench mix (1.5 M NaAc, pH 5.2, 80 mM EDTA, 6 M β-mercaptoethanol), samples were subjected to phenol–chloroform extraction and precipitated with ethanol. Pellets were dissolved in 20 μl of water, incubated with 100 μl of 0.5 M piperidine at 95°C for 20 min. Following ethanol precipitation, DNA was dissolved in 96% formamide.

The changes in reactivity of the accessible (single-stranded or unstacked) T residues in the NT DNA between positions −10 to +20 relative to the transcription start site were evaluated by ImageQuant software. As the shape of the peaks did not vary dramatically, we used their heights to evaluate the relative $KMnO_4$ reactivity at each position. If one assumes that the area below a peak can be roughly approximated by the area of a triangle, two possibilities can be considered: these triangles are (i) isosceles and similar or (ii) isosceles and equal base. A change in the height of the two triangles by a factor of n would correspond to a change in their areas by a factor n^2 or n, respectively. We used the range from n to n^2 to estimate the effect of Tth $NusG^N$ on the accessibility of each residue.

Transcript cleavage

Linear pIA226 DNA template, holo Tth RNAP (200 nM), ApU (100 μM) and starting NTPs (1 μM GTP, 5 μM ATP and UTP, 10 μCi [γ^{32}P]-GTP, 3000 Ci/mmol) were mixed in 30 μl of buffer TGA2 and incubated for 10 min at 55°C. Halted A26 complex was purified by gel filtration through G-50 spin columns equilibrated in TGA2, diluted 4-fold and stored on ice. Reactions were initiated by shifting samples to 55°C. PP_i (1/10 volume of 2 mM stock) and Tth NusG where added where indicated. Samples were quenched with the STOP buffer at the selected times.

Sample analysis

Samples were heated for 2–3 min at 95°C and separated by electrophoresis in denaturing 5–12% acrylamide (19:1)

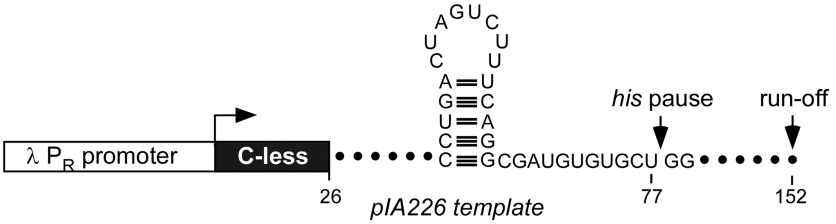
gels (7 M Urea, 0.5× TBE). The gels were dried and RNA products were visualized and quantified using a Molecular Dynamics Storm 820 Phosphorimaging System, ImageQuant Software and Microsoft Excel.

RESULTS

Tth NusG slows down the already 'fast' Tth RNAP

We wanted to test if Tth NusG increases the rate of RNA chain extension *in vitro*. Although it is the most readily observable *in vitro* phenotype of Eco NusG (12,33,34), several considerations suggest that it may not represent the universal effect of NusG-like factors. First, we argued that RfaH (or the isolated $RfaH^N$) acts by disfavoring isomerization into a paused state: (i) $RfaH^N$ decreases the elongation rate on templates that do not encode strong pauses and (ii) the 'fast' RNAP variants that do not pause readily are relatively resistant to the RfaH action (27). Second, similarly to Bsu enzyme, which does not pause at several *E. coli* pause signals (35), Tth RNAP is missing the β′ SI3 domain; deletion of this domain makes Eco RNAP 'fast' (27). If Tth enzyme was resistant to falling into a paused state, it would fail to respond to NusG. Third, Bsu NusG increases pausing at a hairpin-dependent site (14).

It is currently unknown whether regulatory pause signals exist in *Thermus*. We first compared recognition of a canonical hairpin-dependent *his* pause signal by the Eco and Tth enzymes. We used a template that has a phage λ P_R promoter, which is readily recognized by both enzymes (36), followed by a 26-nt long region that does not encode C residues (Figure 2). When transcription is initiated in the absence of CTP, RNAP stalls after addition of AMP at position 26. Upon addition of a complete set of NTPs and heparin (to prevent re-initiation), a single round of elongation can be monitored. Because transcription initiation is quite sensitive to the reaction temperature, we formed halted radiolabeled A26 TECs at near-physiological conditions, 37°C and



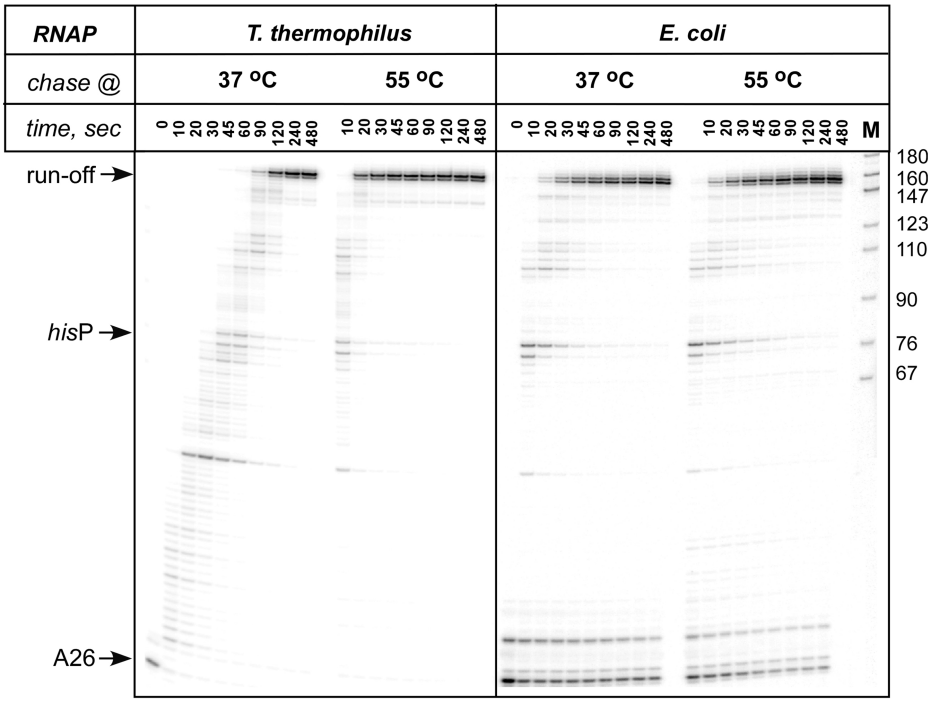


Figure 2. Effects of temperature on transcript elongation of Eco and Tth RNAPs. (Top) Transcript generated from the λ P_R promoter on a linear pIA226 DNA template; transcription start site (a bent arrow), C-less region, *his*P pause signal (pause after U77), and transcript end are indicated on top. (Bottom) Halted A26 TECs were formed at 50 nM with Tth RNAP (left panel) or Eco RNAP (right panel) as described in 'Materials and Methods' section. Elongation was restarted upon addition of NTPs and heparin at 37 or 55°C. Aliquots were withdrawn at times indicated above each lane (in seconds) and analyzed on 8% denaturing gels. Positions of the halted (A26), paused (*his*P) and run-off transcripts are indicated with arrows. Sizes of the ^{32}P-labeled DNA markers used as molecular weight standards (M; pBR322 MspI digest) are indicated on the right.

55°C, for Eco and Tth RNAPs, respectively. We then monitored RNA chain extension upon addition of NTP substrates and heparin; in this case, different temperatures can be used since TECs retain their activity over a broader range of conditions.

We found that Eco RNAP paused at the *hisP* site and elongated the nascent RNA at roughly the same rate at either temperature. Similarly to *T. aquaticus* enzyme, Tth RNAP was 'cold-sensitive' (37) and its elongation rate was dramatically slower at 37°C; however, it still did not recognize the *hisP* signal efficiently (Figure 2). Pausing at the *hisP* site has been proposed to depend in part on interactions between a pause RNA hairpin and the β flap domain of Eco RNAP (38); a failure of Tth RNAP to pause at this site may be due to its inability to establish these contacts.

We next tested the response of Eco and Tth RNAP to both NusG proteins on a template that encodes a hairpin-independent (class II) pause signal, *opsP* (Figure 3A). This signal induces backtracking, a process that is relatively independent of specific interactions between RNAP and the nucleic acid chains in the TEC (12), and thus could be recognized by different RNAPs. While Eco RNAP paused at the *ops* site efficiently at both 37°C and 55°C (left panel), Tth RNAP did not recognize this signal (right panel) even at 37°C; in contrast,

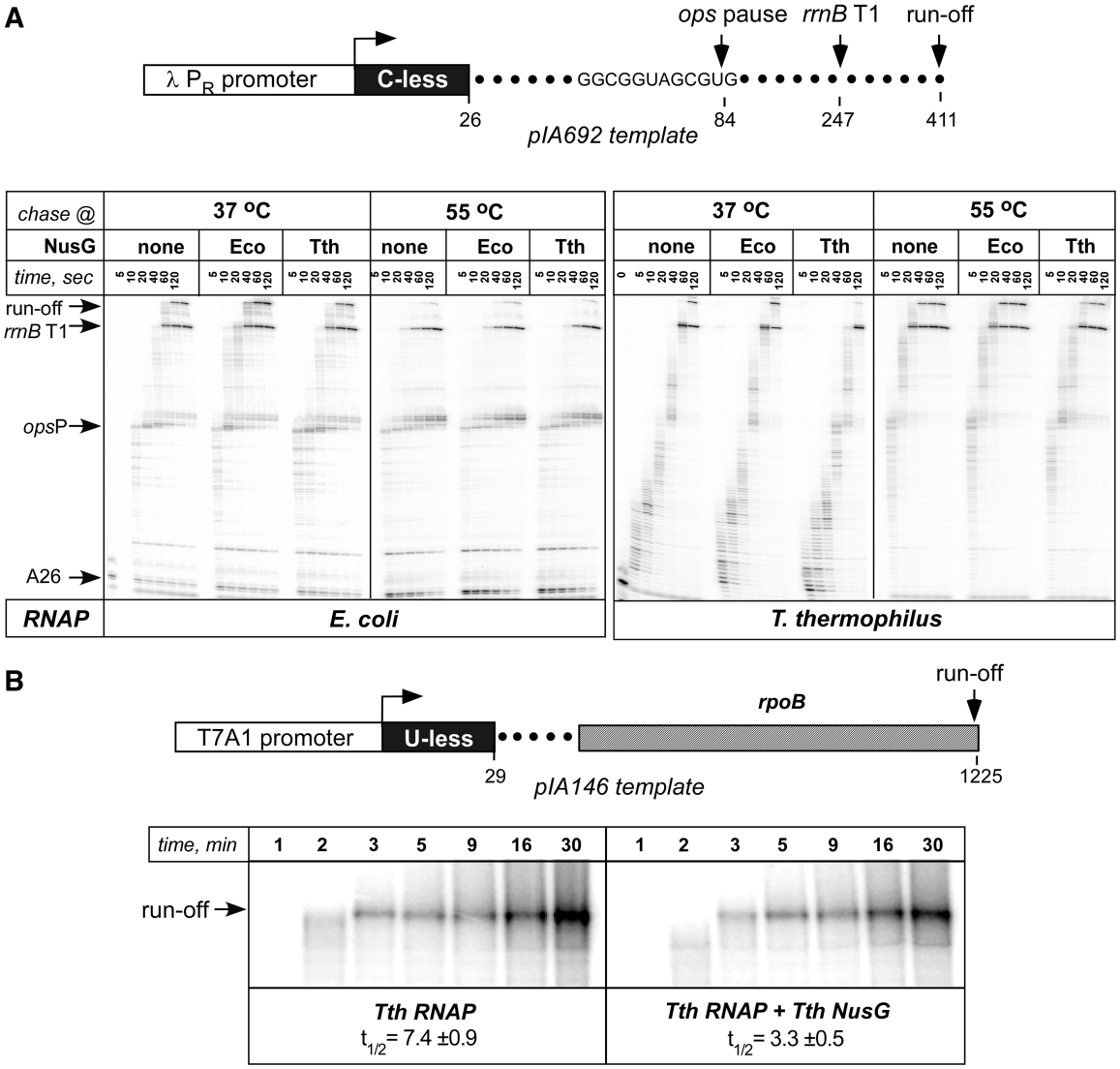


Figure 3. Effects of the NusG proteins on the transcript elongation rate. (**A**) Transcript generated from the λ P_R promoter on a linear pIA692 DNA template; transcription start site (a bent arrow), C-less region, *ops*P pause site (+84), *rrnB* T1 terminator (release site at +247), and transcript end are indicated. (Bottom) Halted A26 TECs were formed at 50 nM with Eco RNAP (left panel) or Tth RNAP (right panel) as described in 'Materials and Methods' section. Elongation was restarted upon addition of NTPs and heparin in the absence or in the presence of 1 μM NusG variant (as indicated above each panel); assays were carried out at 37°C (left panels) or 55°C (right panels). Aliquots were withdrawn at times indicated above each lane and analyzed on 8% denaturing gels. Positions of the halted (A26), paused (*ops*P), terminated (*rrnB* T1) and run-off transcripts are indicated with arrows. (**B**) Effect of the Tth NusG on transcription on a 'pause-free' pIA146 template (top). Halted radiolabeled A29 TECs were extended with NTPs in the absence or in the presence of Tth NusG, aliquots were withdrawn at indicated times and analyzed on 5% denaturing gels; only the gel region near the run-off transcript is shown. The mean rate was calculated from three independent experiments as described previously (27).

Bsu enzyme pauses at *ops*, albeit weakly (35). Addition of Eco NusG moderately increased the elongation rate by 20%, consistent with many published reports [(39) and references therein] of the Eco enzyme, but had no effect on Tth RNAP. Conversely, Tth NusG reduced (by ~60%) an overall elongation rate of its cognate enzyme while having no effect on Eco RNAP. We could not detect any significant site-specific effects of Tth NusG (e.g. an appearance of a new pause species) on either pIA692 (Figure 3A) or pIA226 template (data not shown).

To directly compare the effects of Tth and Eco NusGs on the elongation rate, we utilized a well-characterized pIA146 template (Figure 3B) that encodes the *E. coli rpoB* fragment which is devoid of strong pauses. We monitored the overall elongation rate by accumulation of 1225 nt run-off RNA. The Eco NusG effects on this template have been measured both in bulk (9) and in single-molecule (26) experiments; NusG conferred a moderate (10–20%) rate increase under a variety of conditions, including at near-physiological NTP concentrations. However, even under conditions that favor RNAP pausing (low [GTP]), Tth NusG reduced the mean rate ~2-fold (Figure 3B).

Different NusG proteins have small and distinct effects on intrinsic termination

Next, we wanted to find out whether NusG proteins also differ in their effects on intrinsic termination. We used a set of seven templates (A through G) in which λ P_R promoter is followed by a terminator (Figure 4); five of these signals (C–G) induce termination by Eco RNAP. Additional two candidates (A and B) were cloned from the *T. thermophilus* genome (NC_006461; KEGG); both have a canonical terminator structure with a hairpin

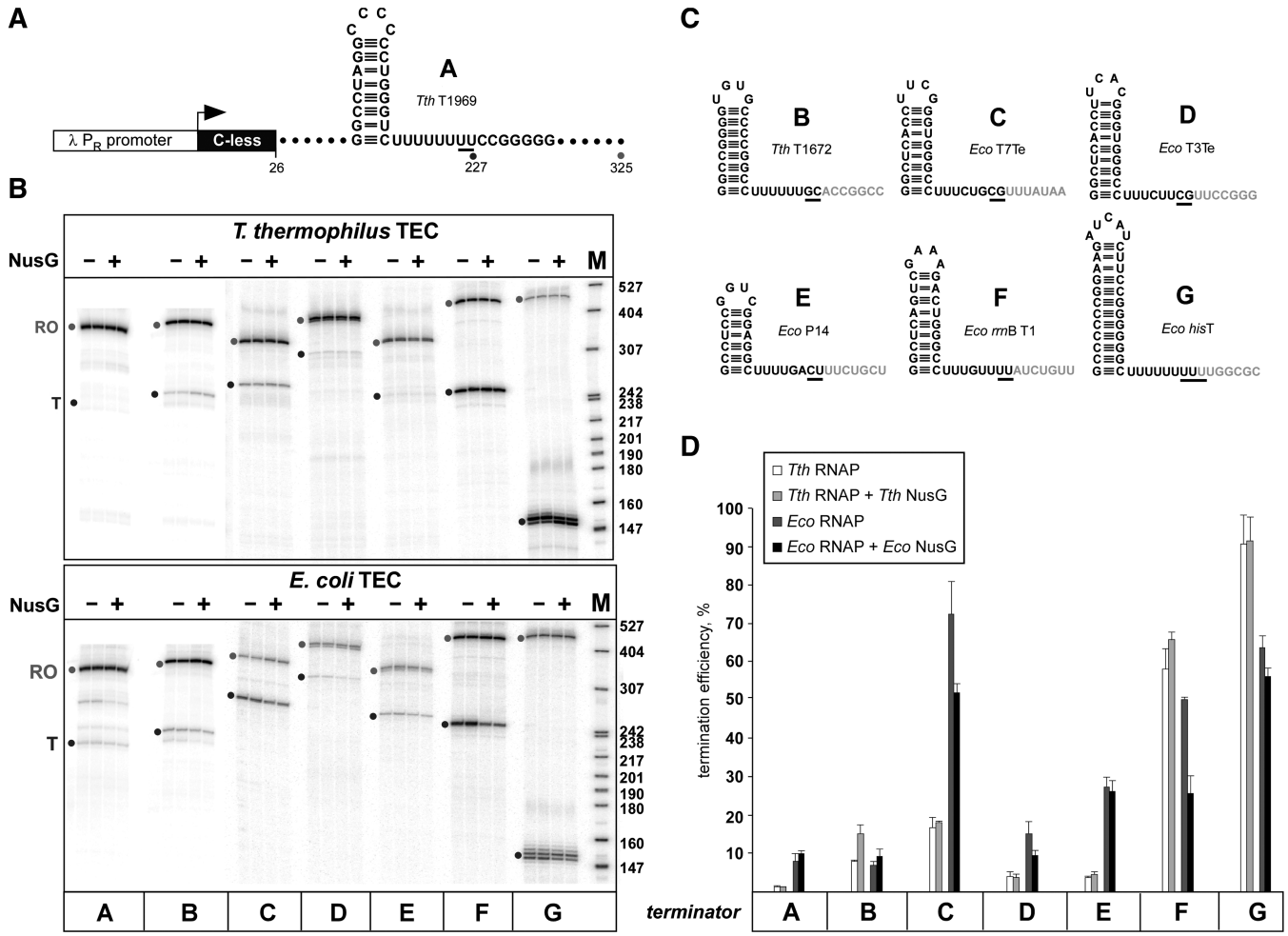


Figure 4. Transcription termination by Eco and Tth RNAPs. (**A**) Transcript generated from the λ P_R promoter on a linear pIA747 DNA (template A); transcription start site (a bent arrow), C-less region (residues 1–26), Tth T_{1969} terminator (release at 227, a red dot) and transcript end (325, a green dot) are indicated. (**B**) Halted A26 TECs were formed at 50 nM on templates A–G (indicated below) with Eco or Tth RNAP. Termination was assayed in single-round A26 RNA extension by addition of all four NTPs (to 200 μM) and heparin (at 10 μg/ml) in the absence or in the presence of 1 μM NusG. The reactions were incubated for 10 min at 55°C (for Tth RNAP, top panel) or at 37°C (for Eco RNAP, bottom panel), and quenched. Products were analyzed on 6% denaturing gels. Positions of terminated (red dots) and run-off (green dots) RNAs are shown on the left. Sizes of the ^{32}P-labeled DNA markers used as molecular weight standards (M; pBR322 MspI digest) are indicated on the right. (**C**) Terminators used in this study. Release occurs between the underlined positions (at two sites in case of *hisT*). (**D**) Termination efficiency (terminated transcript as a fraction of total RNA) was determined in four independent experiments. Templates are indicated below each set of bars, the key is shown in the figure.

followed by a U-rich region and are located at the end of the transcription units. T1672 (template A) is located after the gene encoding isocitrate lyase (genome coordinates 1 409 340–1 409 359); T1969 (template B) is found at the end of the tRNA-Asn gene (1 653 421–1 653 440).

Single round termination assays were carried out at 37°C and 55°C for the Eco and Tth enzymes, respectively (Figure 4B). The two enzymes behaved very differently. While two strong *E. coli* terminators, *his*T (G) and *rrnB* T1 (F), were recognized even better by the Tth RNAP, this enzyme terminated much less efficiently at another strong (T7 Te; C) and two weak, T3 Te (D) and P14 (E), *E. coli* terminators. Two putative *T. thermophilus* terminators triggered weak termination by Eco RNAP; the Tth enzyme essentially bypassed T1969 (A) and terminated with low (and comparable to that of Eco RNAP) efficiency at T1672 (B). While a thorough comparison is impossible to make from such a small data set (and such a comparison was not intended in this work), some preliminary conclusions may be made. At two signals (F and G) that trigger the high-efficiency (>55%) termination by Tth RNAP, the four terminal bases in the transcript are Us. By contrast, at four out of five signals that Tth enzyme bypasses easily (B, C, D and E; <20% termination), the residues at which termination occurs (as measured by the transcript release from the immobilized TECs; IA, data not shown) are GC, CG, CG and CU, respectively. The terminal U residues are thought to favor fraying of the transcript 3′-end from the template DNA during the formation of the elemental pause state, which is a precursor for termination. As compared to the Eco enzyme, Tth RNAP is characterized by a higher apparent rate and reduced pausing (Figures 2 and 3), suggesting that it may be relatively resistant to fraying of the 3′-nucleotide, an effect that would be augmented by the presence of non-U residues at the end of the nascent RNA. The putative Tth 1969 signal appears to be an exception; it induced <10% termination by both enzymes despite the presence of a perfect run of eight U residues at the end of the terminated RNA. In this case, one may consider a possibility that the predicted structure does not form during transcription. For example, an alternative stable RNA structure may form upstream of the putative terminator hairpin, thereby precluding folding of the latter; we did not assess this possibility experimentally.

Response to NusG also differed between the two enzymes. While the effects were generally modest, Eco NusG reproducibly reduced termination at strong *rrnB* T1 and T7 Te sites, and showed small effects at T3 Te and *his*T signals. In contrast, Tth NusG did not reduce termination at any site, and slightly increased termination at *rrnB* T1 and Tth T1672 signals. We did not test the NusG effects on heterologous RNAPs because we did not observe any such effects during elongation assays (Figure 3).

We conclude that (i) NusG proteins from *E. coli* and *T. thermophilus* have modest effects on termination by their cognate RNAPs (Figure 4), and (ii) these effects parallel those observed during transcript elongation (Figure 3).

Tth NusG reduces σ^A-induced pausing by Tth RNAP

Interactions of the *E. coli* initiation factor σ^{70} with the NT DNA strand and the β′ CH domain apparently induce RNAP pausing at promoter-proximal and downstream sequences (40–42) that bear resemblance to the TATAA T hexamer. Pausing is likely caused by energetically favorable σ/DNA contacts that have to be broken before RNAP can move forward. In the RfaH/TEC model (23), $RfaH^N$ domain simultaneously binds to the NT DNA strand and the β′ CH domain, the same targets that σ uses for recruitment to the TEC. We showed that the full-length Eco RfaH and the isolated $RfaH^N$ abrogate σ-dependent pausing of Eco RNAP, presumably through steric competition (43). Since both sets of interactions required for the σ-induced pausing are conserved between the Eco and Tth RNAPs, we reasoned that σ^A, the primary σ factor from *T. thermophilus,* should induce pausing at the TATAAT consensus element. Furthermore, the high degree of conservation of the $NusG^N$-like domains (Figure 1) suggests that Tth NusG may also inhibit σ-dependent pausing.

To test these predictions, we constructed a template with a consensus extended −10 (TGcTATAAT) sequence located downstream from the λ P_R promoter and the C-less region (Figure 5A). We prepared halted A26 TECs and monitored RNA chain extension upon addition of NTPs. Addition of σA to 500 nM induced Tth RNAP pausing downstream from the −10 hexamer (Figure 5B). Since this concentration is far below the cellular level of the *E. coli* σ^{70} (44), this effect is likely physiologically relevant; σ^A concentration is expected to fall in the same range. Both the site and the efficiency of σ^A-induced pausing (Figure 5B) were similar to those observed under identical experimental conditions for the σ^{70}-induced pausing of Eco RNAP (43). This finding is not surprising: primary σ factors recognize the −10 element, and the determinants that ensure specific promoter recognition (the bases on the NT strand and the β′ CH residues) are also accessible within the TEC to mediate σ-induced stalling at promoter-like sites during elongation (42). When present at 500 nM, Tth NusG reduced the fraction of σ^A-stalled TECs by ~3-fold (Figure 5B). Given that Tth NusG decreases the overall rate of elongation (Figure 3), its effect at the σ-dependent site cannot be explained by the Eco NusG-like anti-pausing activity and is most likely due to steric competition. Interestingly, in contrast to $RfaH^N$ (23), Tth $NusG^N$ was less effective in reducing σ^A-induced pausing (data not shown), suggesting that is has a lower affinity to the TEC than the full-length protein; the same effect was reported for Eco $NusG^N$ (9).

Tth NusG binds to NT DNA in the TEC

In the RfaH/TEC model (27), $RfaH^N$ interacts with the upstream fork junction where the NT DNA bends sharply (~90°) to reanneal with the template strand. A highly similar $NusG^N$ is expected to bind to the same target. To examine whether the fork-junction accessibility is affected by Tth $NusG^N$, we used footprinting by $KMnO_4$, which reacts with unstacked or unpaired

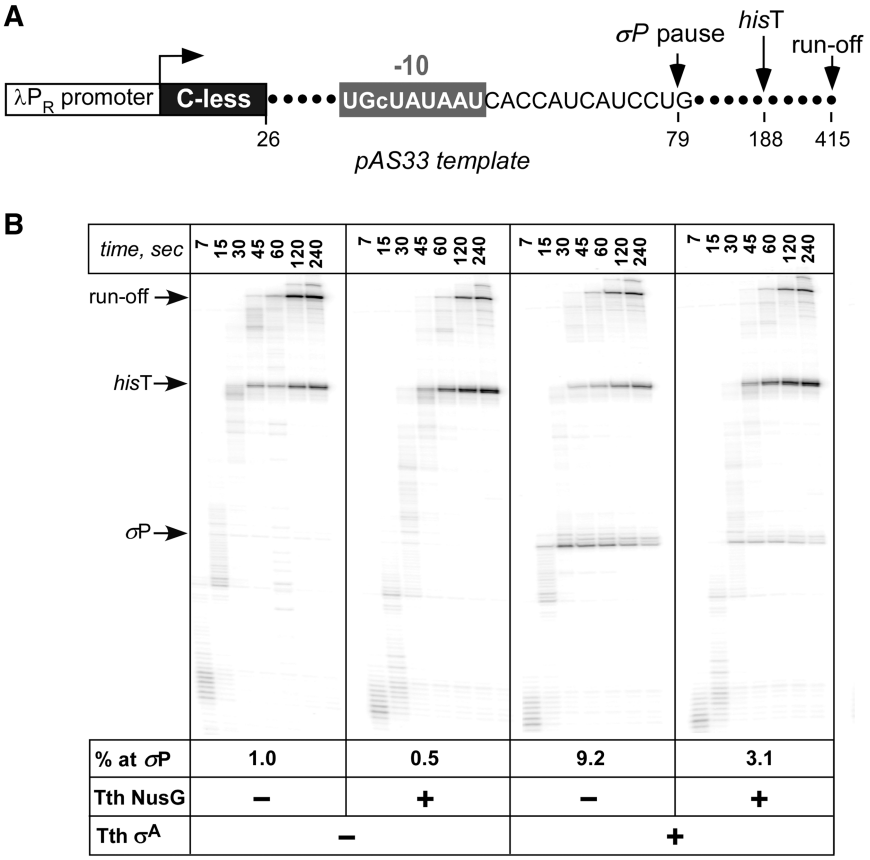


Figure 5. Tth NusG inhibits σ^A-induced pausing by Tth RNAP. (**A**) The linear DNA template is shown on top with the transcription start site (a bent arrow), the extended −10 motif, and the end indicated. (**B**) Single-round pause assays were performed in the absence or in the presence of σ^A and Tth NusG (at 500 nM each), where indicated. A representative 6% denaturing gel is shown. Position of σP pause site was mapped in the presence of chain-terminating NTPs (data not shown). The fraction of RNA at the σP site after a 240-s incubation (as percentage of total RNA) is presented below each panel.

thymines (Figure 6A). We chose to use the isolated $NusG^N$ domain in place of the full-length protein because (i) we expected that the presence of a flexibly tethered $NusG^C$ (18) could complicate structural studies and (ii) the N-terminal domains of *E. coli* RfaH and NusG are sufficient for their effects on elongation (9,23).

The structure of Tth TEC (45) suggests that in halted A26 complexes the fork junction lies between +14 and +15 residues, and that the NT stand T residues at +15, +16, +18 and +20 are single stranded (Figure 6A, TEC). For comparison, in open promoter complexes (Figure 6A, RP_o) the NT strand T residues at −10, −7, −4 and −3 are expected to be single-stranded and +2—unstacked, assuming that σ^A and σ^{70} make similar contacts to the −10 element. However, σ^{70} contacts with the −10 and −7 bases block $KMnO_4$ access to the plane of these bases, which thereby appear highly resistant to modification (36). Indeed, the expected patterns were observed in both RP_o and TEC (Figure 6B).

We found that addition of Tth $NusG^N$ to the halted TEC led to a partial protection of the upstream part of the bubble (see 'Materials and Methods' section). In the presence of Tth $NusG^N$, the permanganate reactivity at positions +15 and +16 was reduced by ~40–60% and 30–50%, respectively. By contrast, accessibility of the T residues at +18 (<15% change) and +20 (no detectable change) was not affected. Bsu NusG has also been shown to alter $KMnO_4$ reactivity of the NT strand *in vivo* and *in vitro* (14).

Tth NusG stabilizes the post-translocated state of the TEC

RfaH binds to a similar position on the nucleic acids within the TEC and favors forward translocation by the Eco RNAP (27). We argued that this activity may explain an observed decrease in pausing conferred by RfaH. A recent single-molecule study (26) has demonstrated that Eco NusG increases Eco RNAP velocity and decreases the entry into backtracked paused states. Both effects can be readily explained by an effect on translocation; indeed, the same result was achieved by applying assisting force. To test whether this is a general feature of the NusG-like proteins, we examined the effect of Tth NusG on two types of RNA cleavage reactions. Pyrophosphorolysis, a reversal of the nucleotide addition reaction, and intrinsic RNA hydrolysis occur in pre-translocated TECs; sensitivity of the nascent RNA to these reactions can be used to infer the position of the nascent RNA in the active site (Figure 6C).

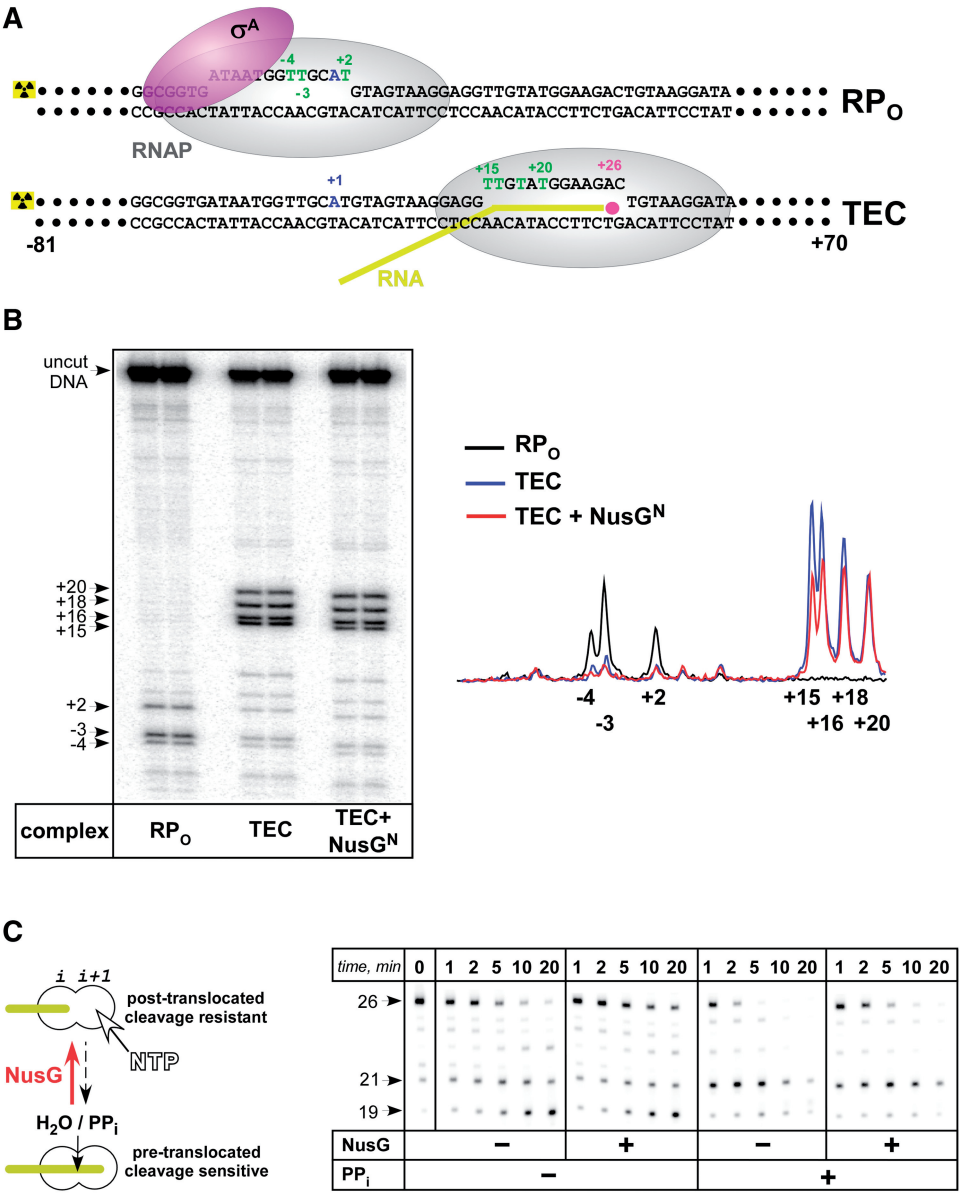


Figure 6. Tth NusG binds to Tth TEC and favors forward translocation. (**A**) Transcription complexes were formed on the λ P_R promoter template; the top (NT) strand is labeled. The start site (+1) is indicated in blue. The NT strand T residues known (36) or expected to be modified by $KMnO_4$ are shown in green. In RP_o, the T residues at −10 and −7 positions are expected to be protected by the bound σ (magenta). In the TEC halted at A26, the Ts at positions +15, +16, +18 and +20 are expected to be accessible. (**B**) Analysis of $KMnO_4$ modification patterns of RP_o and TECs formed with and without Tth $NusG^N$ ('Materials and Methods' section). (Left) Positions of hypersensitive (in one of the complexes) T residues and the full-length, uncut DNA are indicated with arrows. (Right) Trace analysis of the data. The low-level (but above the background) modification of Ts at −4, −3, +2, +4 and +7 positions in elongation complexes is likely due to the persistence of open complexes and abortive complexes. (**C**) Tth NusG effect on RNA cleavage. (Left) TECs may interconvert between states in which the 3′-end of the nascent RNA occupies different positions in the RNAP active site. In the post-translocated configuration, the 3′-OH is in the *i* site, the *i+1* site is poised to bind the incoming NTP, and the complex is resistant to cleavage. In the pre-translocated state, the 3′-end occupies the *i+1* site and the nascent RNA is sensitive to PP_i- and H_2O-mediated cleavage. (Right) Halted Tth A26 complexes were incubated at 55°C for the times shown; NusG (200 nM) and PP_i (200 μM) were present where indicated. The reactions were analyzed on a 12% denaturing gel.

The A26 TEC formed by either Eco (27) or Tth RNAP (46) are relatively sensitive to PP_i and GreB, indicating that these complexes are in a pre-translocated and back-tracked states. We formed halted radiolabeled A26 complexes, removed NTPs by gel filtration, and then incubated these complexes for 1–20 min at 55°C; PP_i (200 μM) and NusG (200 nM) were added where indicated (Figure 6C). As expected, A26 complexes were sensitive to the PP_i-induced cleavage, with a half-life of ∼1.3 min at 55°C. Tth, but not Eco (27), RNAP also displayed a high level of intrinsic hydrolytic activity (the left panel). Addition of NusG increased the half-life of A26 approximately 2-fold in both cases, from 3.7 to 7.6 min in the absence of PP_i and from 1.3 to 2.8 min in the presence of PP_i. The two reactions differ in their mechanisms, phosphoryl transfer and water-mediated, transcript-assisted

hydrolysis (47), but are similarly affected by Tth NusG. These results suggest that, similarly to Eco NusG and RfaH, Tth NusG favors forward translocation of RNAP along the DNA.

DISCUSSION

In this work we demonstrate that *T. thermophilus* elongation factor NusG (i) apparently binds to the upstream fork junction of Tth TEC; (ii) stimulates forward translocation of RNAP; and (iii) competes with the initiation factor σ during elongation; these phenotypes and the binding site on RNAP are shared by *E. coli* RfaH and NusG. On the other hand, in contrast to RfaH and NusG which both decrease RNAP pausing, Tth NusG does not facilitate transcript elongation by its cognate RNAP. Thus, it appears that the well-documented acceleration of Eco RNAP by NusG may not be an essential activity of this universally conserved transcription factor. Together with recent reports from other groups, our data suggest that even though all NusG-like proteins likely bind to the same site on TEC (Figure 7), their regulatory outcomes may depend on the intrinsic properties of the affected RNAP and the identity of their interaction partners.

NusG interactions with the TEC

In Eco RfaH (23) and NusG (9), Mja NusG (28) and human DSIF (48), the binding to RNAP is likely mediated by interactions between the CH domain and a hydrophobic patch on the N-terminal domain. Thus, the mode of recruitment to RNAP is likely common within the NusG superfamily. However, the details of interactions with nucleic acids may vary significantly among different proteins. Eco RfaH requires an *ops* sequence in the NT DNA strand exposed on the RNAP surface during recruitment and directly crosslinks to the NT strand in the TEC (19). RfaH likely maintains non-specific interactions with the DNA throughout elongation; RfaH recognizes an *ops* site positioned far downstream from its site of recruitment. In contrast, Eco NusG does not crosslink to DNA (or RNA); consistently, the region that mediates RfaH binding to *ops* is the least conserved between the two proteins (25). Two NusG orthologs, from *Aquifex aeolicus* and *T. maritima*, bind to nucleic acids non-specifically (15,21); this property could be attributed to a large positively charged domains inserted into a flexible β loop in the $NusG^N$ domain (Figure 1). Finally, the action of Bsu NusG action is sensitive to sequence alterations (49), and Eco NusG exhibits differential effects at different pause sites (26). It is, however, unclear whether these effects are mediated by base-specific contacts between NusG and nucleic acids or are conferred by changes in the TEC induced by altered RNA–DNA interactions.

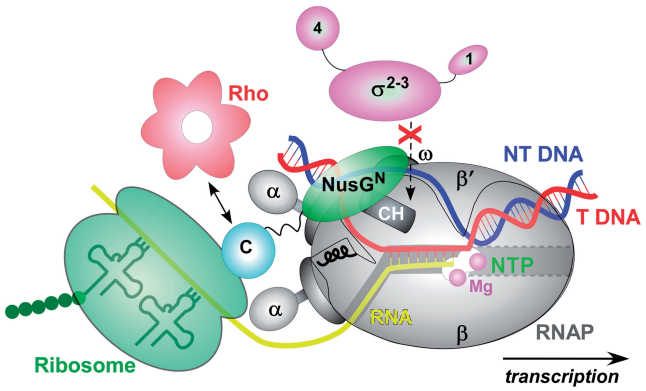


Figure 7. A model for Tth NusG interactions. In the TEC, RNAP (grey) is bound to the template (T, red) and NT (blue) DNA strands that are separated in front of the active site (white circles, shown with two catalytic Mg ions and the substrate NTP) to form a transcription bubble. The single-stranded T strand is paired with the nascent RNA (yellow) in an 8–9 bp RNA:DNA hybrid, the NT strand is exposed on the RNAP surface. $NusG^N$ (green) interacts with the hydrophobic tip of the β′ CH (dark grey cylinder) near the upstream fork junction. $NusG^C$ (cyan) is connected to $NusG^N$ via a flexible linker and may interact with ribosome (as shown), Rho (particularly in the absence of translation), upstream DNA or other regulators. In NusG-bound TEC, the primary binding site for σ (magenta) composed of the NT strand nucleotides and the β′ CH is blocked.

The observed protection of the residues at the fork-junction by Tth NusG (Figure 6B) could be interpreted in several ways. First, NusG could make contacts to the DNA bases that would directly shield Ts from $KMnO_4$ attack. Second, NusG could induce changes in stacking interaction or backbone distortion that indirectly result in altered sensitivity to $KMnO_4$. Third, as suggested earlier (10,12) and recently supported by the single-molecule data for Eco NusG (26), Tth NusG could favor forward translocation of RNAP. In the latter case, one base pair would reanneal at the upstream part of the bubble, leading to protection of T+15 against $KMnO_4$, and one base pair will become separated ahead of the active site. The +2 base (T28) is predicted to be unstacked but may or may not become sensitive to modification by $KMnO_4$; in different transcription complexes, +2 position may appear accessible or protected.

The observed pattern is inconsistent with base-specific contacts between NusG and the NT DNA; such contacts would be expected to confer strong protection (e.g. of the −10 and −7 T residues by the bound σ in RP_o; Figure 6B). Competition with σ (Figure 5) suggests that Tth NusG may sterically hinder $KMnO_4$ access to the bases; however, additional effects on the DNA structure cannot be excluded. This pattern of protection, together with the reduction in RNA cleavage rates (Figure 6C), is consistent with forward translocation induced by NusG bound to the upstream fork junction.

The NusG/NT DNA interactions may (i) serve as specific signals for a regulator (e.g. RfaH) recruitment; (ii) insulate the TEC from factors (e.g. σ) that target the NT DNA; (iii) constrain the path/conformation of the NT strand on the RNAP surface; and (iv) help to stabilize the fork junction structure; *A. aeolicus* NusG was proposed to induce partial melting of duplex DNA (21). However, at least in some cases the NT contacts appear to be dispensable; for example, archaeal NusG functions in the absence of the NT DNA (28).

NusG role in transcriptional pausing

We propose that the NusG-like proteins play a dual role in regulation of transcriptional pausing: they insulate any elongating RNAP from re-binding of σ (and other factors that target the NT DNA or the β′ CH) and prevent isomerization of some, pause-prone, RNAP molecules into the paused state.

The ability of Eco NusG to increase the elongation rate was noted early and proposed to be an important part of its mechanism. Similarly, RfaH is thought to act by decreasing RNAP pausing. Pausing is triggered by signals that induce formation of an elemental pause state, in which the nascent RNA 3′-end is misaligned (50,51); isomerization into the paused state is thought to occur from the pre-translocated state (50). RfaH fails to affect transcription on pause-free templates and with pause-resistant RNAPs (38) and Eco NusG has only a small effect on pause-free RNAP velocity (26). Both RfaH and NusG appear to favor the forward-translocated state of the TEC, thereby inhibiting isomerization into the paused state (10,27).

Recent data suggest, however, that the antipausing activity is not shared by all bacterial NusGs: Tth (Figure 2) and Bsu (14) factors increase rather than decrease pausing. These differences likely reflect the properties of the RNAP species: while Eco RNAP pauses frequently, at least during transcription *in vitro*, Bsu (35) and Tth enzyme (Figures 2 and 3) do not recognize pause signals that hinder Eco RNAP even when moving at approximately the same rate. Pyrophosphorolysis in static TECs (e.g. A26) reveals that the nascent RNA in Tth TECs is significantly more resistant to cleavage as compared to the Eco RNAP [the half-life of 13 s at 25 μM PP_i (27)], suggesting that Tth RNAP may be stabilized in the post-translocated state. If this were true, Tth RNAP would be expected to ignore pauses and transcribe at a faster rate, as seen in bulk elongation assays (Figures 2 and 3). These observations suggest that translocation may not be rate limiting for Tth RNAP; in this case, no 'stimulatory' effect of NusG on elongation would be expected.

RNAPs that are not accelerated by NusG (or RfaH) share one common property—they transcribe at an overall faster rate in the absence of accessory factors. This could be due to sequence-specific differences between the catalytic elements which confer many enzyme-specific properties (37). While it is difficult to point out a single underlying reason, our attempts to convert Eco RNAP into a Bsu-like enzyme through removal of large *E. coli*-specific insertions have demonstrated that the β′ SI3 domain is responsible for many difference observed between Eco and Bsu enzymes *in vitro* (52). Eco ΔSI3 enzyme transcribes at a faster rate, pauses and terminates less efficiently, and is slowed down by $RfaH^N$ (27). Similarly, Bsu (14) and Tth (Figure 2) NusGs reduce the elongation rate by their cognate RNAPs which are already inherently fast. It is currently unclear how Bsu and Tth NusGs slow RNAP down; the single-molecule analysis would be required to distinguish between the effects on pausing (isomerization into, or escape from, the off-pathway states) and pause-free elongation between the short-lived pauses. The available data appear to suggest that one of the roles of NusG could be to maintain the certain rate of RNA synthesis rather than to speed the RNAP up. In bacteria, this requirement may be imposed by the need to couple transcription to translation (53).

What is the main role of NusG in the cell?

NusG-like proteins are present in all three kingdoms of life and have been implicated in various essential cellular processes. Eco NusG reduces RNAP pausing and intrinsic termination (4), recruits Rho to the TEC (11,13) and participates in formation of multi-protein complexes (9) that mediate antitermination modification of RNAP transcribing *rrn* and phage λ genes and assist termination by HK022 Nun protein. Recent reports (29,54) suggest that Eco NusG may bind to ribosome; we have proposed a similar role for RfaH (25). Due to this functional diversity, it is difficult to point out which of these activities is responsible for the ubiquity of NusG. In *E. coli,* transcriptional repression of foreign DNA through direct interactions with Rho has been proposed to constitute the essential role of NusG (8). However, both NusG and Rho are dispensable in *B. subtilis* (16) and Tth NusG does not interact with Rho (18).

Binding to RNAP and modulation of its rate are mediated by $NusG^N$ whereas $NusG^C$ domain is thought to establish contacts with other partners (Figure 7). We hypothesize that the primary role of $NusG^N$ may be to tether $NusG^C$ to the TEC, whereas $NusG^C$ is essential for the assembly and function of the regulatory complexes that include one (Rho) or many (other Nus factors, *nut* RNA, etc.) components.

We also propose that the roles of these domains are conserved in all kingdoms. In Archaea and eukaryotes, NusG homolog Spt5 forms a heterodimer with Spt4; the Spt4/5 complex (called DSIF in humans) enhances RNAP II processivity (55). The N-terminal domain of Mja NusG likely binds to the same site on RNAP (28). The C-terminal domains contain several KOW motifs (as compared to just one in NusG) and repeats, which, upon phosphorylation by cyclin-dependent kinases (P-TEFb in humans and Bur in yeast), appear to nucleate assembly of large protein complexes with diverse regulatory functions. C-terminal repeats promote RNAPII elongation, recruitment of the PAF complex and mRNA capping enzymes, histone H2B K123 monoubiquitination and histone H3 K4 and K36 trimethylation, suppress Rad26-independent transcription-coupled nucleotide excision repair (30,55,56), and may couple transcription by RNAP I to rRNA processing and ribosome assembly (57).

ACKNOWLEDGEMENTS

The authors thank Vladimir Svetlov for the construction of the Tth NusG expression plasmid and many stimulating discussions.

FUNDING

National Institutes of Health (GM67153).

Conflict of interest statement. None declared.

REFERENCES

1. Li,J., Horwitz,R., McCracken,S. and Greenblatt,J. (1992) NusG, a new Escherichia coli elongation factor involved in transcriptional antitermination by the N protein of phage lambda. *J. Biol. Chem.*, **267**, 6012–6019.
2. Sullivan,S.L. and Gottesman,M.E. (1992) Requirement for E. coli NusG protein in factor-dependent transcription termination. *Cell*, **68**, 989–994.
3. Hung,S.C. and Gottesman,M.E. (1995) Phage HK022 Nun protein arrests transcription on phage lambda DNA in vitro and competes with the phage lambda N antitermination protein. *J. Mol. Biol.*, **247**, 428–442.
4. Burova,E., Hung,S.C., Sagitov,V., Stitt,B.L. and Gottesman,M.E. (1995) Escherichia coli NusG protein stimulates transcription elongation rates in vivo and in vitro. *J. Bacteriol.*, **177**, 1388–1392.
5. Zellars,M. and Squires,C.L. (1999) Antiterminator-dependent modulation of transcription elongation rates by NusB and NusG. *Mol. Microbiol.*, **32**, 1296–1304.
6. Squires,C.L., Greenblatt,J., Li,J., Condon,C. and Squires,C.L. (1993) Ribosomal RNA antitermination in vitro: requirement for Nus factors and one or more unidentified cellular components. *Proc. Natl Acad. Sci. USA*, **90**, 970–974.
7. Mooney,R.A., Davis,S.E., Peters,J.M., Rowland,J.L., Ansari,A.Z. and Landick,R. (2009) Regulator trafficking on bacterial transcription units in vivo. *Mol. Cell*, **33**, 97–108.
8. Cardinale,C.J., Washburn,R.S., Tadigotla,V.R., Brown,L.M., Gottesman,M.E. and Nudler,E. (2008) Termination factor Rho and its cofactors NusA and NusG silence foreign DNA in E. coli. *Science*, **320**, 935–938.
9. Mooney,R.A., Schweimer,K., Rosch,P., Gottesman,M. and Landick,R. (2009) Two structurally independent domains of E. coli NusG create regulatory plasticity via distinct interactions with RNA polymerase and regulators. *J. Mol. Biol.*, **391**, 341–358.
10. Pasman,Z. and von Hippel,P.H. (2000) Regulation of rho-dependent transcription termination by NusG is specific to the Escherichia coli elongation complex. *Biochemistry*, **39**, 5573–5585.
11. Burns,C.M. and Richardson,J.P. (1995) NusG is required to overcome a kinetic limitation to Rho function at an intragenic terminator. *Proc. Natl Acad. Sci. USA*, **92**, 4738–4742.
12. Artsimovitch,I. and Landick,R. (2000) Pausing by bacterial RNA polymerase is mediated by mechanistically distinct classes of signals. *Proc. Natl Acad. Sci. USA*, **97**, 7090–7095.
13. Li,J., Mason,S.W. and Greenblatt,J. (1993) Elongation factor NusG interacts with termination factor rho to regulate termination and antitermination of transcription. *Genes Dev.*, **7**, 161–172.
14. Yakhnin,A.V., Yakhnin,H. and Babitzke,P. (2008) Function of the Bacillus subtilis transcription elongation factor NusG in hairpin-dependent RNA polymerase pausing in the trp leader. *Proc. Natl Acad. Sci. USA*, **105**, 16131–16136.
15. Liao,D., Lurz,R., Dobrinski,B. and Dennis,P.P. (1996) A NusG-like protein from Thermotoga maritima binds to DNA and RNA. *J. Bacteriol.*, **178**, 4089–4098.
16. Ingham,C.J., Dennis,J. and Furneaux,P.A. (1999) Autogenous regulation of transcription termination factor Rho and the requirement for Nus factors in Bacillus subtilis. *Mol. Microbiol.*, **31**, 651–663.
17. Xia,M., Lunsford,R.D., McDevitt,D. and Iordanescu,S. (1999) Rapid method for the identification of essential genes in Staphylococcus aureus. *Plasmid*, **42**, 144–149.
18. Reay,P., Yamasaki,K., Terada,T., Kuramitsu,S., Shirouzu,M. and Yokoyama,S. (2004) Structural and sequence comparisons arising from the solution structure of the transcription elongation factor NusG from Thermus thermophilus. *Proteins*, **56**, 40–51.
19. Artsimovitch,I. and 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.
20. Chatzidaki-Livanis,M., Coyne,M.J. and Comstock,L.E. (2009) A family of transcriptional antitermination factors necessary for synthesis of the capsular polysaccharides of Bacteroides fragilis. *J. Bacteriol.*, **191**, 7288–7295.
21. Steiner,T., Kaiser,J.T., Marinkovic,S., Huber,R. and Wahl,M.C. (2002) Crystal structures of transcription factor NusG in light of its nucleic acid- and protein-binding activities. *EMBO J.*, **21**, 4641–4653.
22. Knowlton,J.R., Bubunenko,M., Andrykovitch,M., Guo,W., Routzahn,K.M., Waugh,D.S., Court,D.L. and Ji,X. (2003) A spring-loaded state of NusG in its functional cycle is suggested by X-ray crystallography and supported by site-directed mutants. *Biochemistry*, **42**, 2275–2281.
23. Belogurov,G.A., Vassylyeva,M.N., Svetlov,V., Klyuyev,S., Grishin,N.V., Vassylyev,D.G. and Artsimovitch,I. (2007) Structural basis for converting a general transcription factor into an operon-specific virulence regulator. *Mol. Cell*, **26**, 117–129.
24. Zhou,H., Liu,Q., Gao,Y., Teng,M. and Niu,L. (2009) Crystal structure of NusG N-terminal (NGN) domain from Methanocaldococcus jannaschii and its interaction with rpoE''. *Proteins*, **76**, 787–793.
25. Belogurov,G.A., Mooney,R.A., Svetlov,V., Landick,R. and Artsimovitch,I. (2009) Functional specialization of transcription elongation factors. *EMBO J.*, **28**, 112–122.
26. Herbert,K.M., Zhou,J., Mooney,R.A., Porta,A.L., Landick,R. and Block,S.M. (2010) E. coli NusG inhibits backtracking and accelerates pause-free transcription by promoting forward translocation of RNA polymerase. *J. Mol. Biol.*, **399**, 17–30.
27. Svetlov,V., Belogurov,G.A., Shabrova,E., Vassylyev,D.G. and Artsimovitch,I. (2007) Allosteric control of the RNA polymerase by the elongation factor RfaH. *Nucleic Acids Res.*, **35**, 5694–5705.
28. Hirtreiter,A., Damsma,G.E., Cheung,A.C., Klose,D., Grohmann,D., Vojnic,E., Martin,A.C., Cramer,P. and Werner,F. (2010) Spt4/5 stimulates transcription elongation through the RNA polymerase clamp coiled-coil motif. *Nucleic Acids Res.*, **38**, 4040–4051.
29. Burmann,B.M., Schweimer,K., Luo,X., Wahl,M.C., Stitt,B.L., Gottesman,M.E. and Rosch,P. (2010) A NusE:NusG complex links transcription and translation. *Science*, **328**, 501–504.
30. Zhou,K., Kuo,W.H., Fillingham,J. and Greenblatt,J.F. (2009) Control of transcriptional elongation and cotranscriptional histone modification by the yeast BUR kinase substrate Spt5. *Proc. Natl Acad. Sci. USA*, **106**, 6956–6961.
31. Belogurov,G.A., Sevostyanova,A., Svetlov,V. and Artsimovitch,I. (2010) Functional regions of the N-terminal domain of the antiterminator RfaH. *Mol. Microbiol.*, **76**, 286–301.
32. Studier,F.W. (2005) Protein production by auto-induction in high density shaking cultures. *Protein Exp. Purification*, **41**, 207–234.
33. Burova,E., Hung,S.C., Chen,J., Court,D.L., Zhou,J.G., Mogilnitskiy,G. and Gottesman,M.E. (1999) Escherichia coli nusG mutations that block transcription termination by coliphage HK022 Nun protein. *Mol. Microbiol.*, **31**, 1783–1793.
34. Burns,C.M., Richardson,L.V. and Richardson,J.P. (1998) Combinatorial effects of NusA and NusG on transcription elongation and Rho-dependent termination in Escherichia coli. *J. Mol. Biol.*, **278**, 307–316.
35. Artsimovitch,I., Svetlov,V., Anthony,L., Burgess,R.R. and Landick,R. (2000) RNA polymerases from Bacillus subtilis and Escherichia coli differ in recognition of regulatory signals in vitro. *J. Bacteriol.*, **182**, 6027–6035.
36. Belogurov,G.A., Vassylyeva,M.N., Sevostyanova,A., Appleman,J.R., Xiang,A.X., Lira,R., Webber,S.E., Klyuyev,S., Nudler,E., Artsimovitch,I. *et al.* (2009) Transcription inactivation through local refolding of the RNA polymerase structure. *Nature*, **457**, 332–335.
37. Miropolskaya,N., Artsimovitch,I., Klimasauskas,S., Nikiforov,V. and Kulbachinskiy,A. (2009) Allosteric control of catalysis by the

F loop of RNA polymerase. *Proc. Natl Acad. Sci. USA*, **106**, 18942–18947.
38. Toulokhonov,I., Artsimovitch,I. and Landick,R. (2001) Allosteric control of RNA polymerase by a site that contacts nascent RNA hairpins. *Science*, **292**, 730–733.
39. Richardson,L.V. and Richardson,J.P. (2005) Identification of a structural element that is essential for two functions of transcription factor NusG. *Biochim. Biophysic. Acta*, **1729**, 135–140.
40. Nickels,B.E., Mukhopadhyay,J., Garrity,S.J., Ebright,R.H. and Hochschild,A. (2004) The sigma 70 subunit of RNA polymerase mediates a promoter-proximal pause at the lac promoter. *Nat. Struct. Mol. Biol.*, **11**, 544–550.
41. Roberts,J.W., Yarnell,W., Bartlett,E., Guo,J., Marr,M., Ko,D.C., Sun,H. and Roberts,C.W. (1998) Antitermination by bacteriophage lambda Q protein. *Cold Spring Harbor Symp. Quant. Biol.*, **63**, 319–325.
42. Mooney,R.A. and Landick,R. (2003) Tethering sigma70 to RNA polymerase reveals high in vivo activity of sigma factors and sigma70-dependent pausing at promoter-distal locations. *Genes Dev.*, **17**, 2839–2851.
43. Sevostyanova,A., Svetlov,V., Vassylyev,D.G. and Artsimovitch,I. (2008) The elongation factor RfaH and the initiation factor sigma bind to the same site on the transcription elongation complex. *Proc. Natl Acad. Sci. USA*, **105**, 865–870.
44. Grigorova,I.L., Phleger,N.J., Mutalik,V.K. and Gross,C.A. (2006) Insights into transcriptional regulation and sigma competition from an equilibrium model of RNA polymerase binding to DNA. *Proc. Natl Acad. Sci. USA*, **103**, 5332–5337.
45. Vassylyev,D.G., Vassylyeva,M.N., Perederina,A., Tahirov,T.H. and Artsimovitch,I. (2007) Structural basis for transcription elongation by bacterial RNA polymerase. *Nature*, **448**, 157–162.
46. Symersky,J., Perederina,A., Vassylyeva,M.N., Svetlov,V., Artsimovitch,I. and Vassylyev,D.G. (2006) Regulation through the RNA polymerase secondary channel. Structural and functional variability of the coiled-coil transcription factors. *J. Biol. Chem.*, **281**, 1309–1312.
47. Zenkin,N., Yuzenkova,Y. and Severinov,K. (2006) Transcript-assisted transcriptional proofreading. *Science*, **313**, 518–520.
48. Wenzel,S., Martins,B.M., Rosch,P. and Wohrl,B.M. (2010) Crystal structure of the human transcription elongation factor DSIF hSpt4 subunit in complex with the hSpt5 dimerization interface. *Biochem. J.*, **425**, 373–380.
49. Yakhnin,A.V. and Babitzke,P. (2010) Mechanism of NusG-stimulated pausing, hairpin-dependent pause site selection and intrinsic termination at overlapping pause and termination sites in the Bacillus subtilis trp leader. *Mol. Microbiol.*, **76**, 690–705.
50. Sydow,J.F., Brueckner,F., Cheung,A.C., Damsma,G.E., Dengl,S., Lehmann,E., Vassylyev,D. and Cramer,P. (2009) Structural basis of transcription: mismatch-specific fidelity mechanisms and paused RNA polymerase II with frayed RNA. *Mol. Cell*, **34**, 710–721.
51. Toulokhonov,I., Zhang,J., Palangat,M. and Landick,R. (2007) A central role of the RNA polymerase trigger loop in active-site rearrangement during transcriptional pausing. *Mol. Cell*, **27**, 406–419.
52. Artsimovitch,I., Svetlov,V., Murakami,K.S. and Landick,R. (2003) Co-overexpression of Escherichia coli RNA polymerase subunits allows isolation and analysis of mutant enzymes lacking lineage-specific sequence insertions. *J. Biol. Chem.*, **278**, 12344–12355.
53. Roberts,J.W. (2010) Molecular biology. Syntheses that stay together. *Science*, **328**, 436–437.
54. Proshkin,S., Rahmouni,A.R., Mironov,A. and Nudler,E. (2010) Cooperation between translating ribosomes and RNA polymerase in transcription elongation. *Science*, **328**, 504–508.
55. Chen,Y., Yamaguchi,Y., Tsugeno,Y., Yamamoto,J., Yamada,T., Nakamura,M., Hisatake,K. and Handa,H. (2009) DSIF, the Paf1 complex, and Tat-SF1 have nonredundant, cooperative roles in RNA polymerase II elongation. *Genes Dev.*, **23**, 2765–2777.
56. Pei,Y. and Shuman,S. (2002) Interactions between fission yeast mRNA capping enzymes and elongation factor Spt5. *J. Biol. Chem.*, **277**, 19639–19648.
57. Schneider,D.A., French,S.L., Osheim,Y.N., Bailey,A.O., Vu,L., Dodd,J., Yates,J.R., Beyer,A.L. and Nomura,M. (2006) RNA polymerase II elongation factors Spt4p and Spt5p play roles in transcription elongation by RNA polymerase I and rRNA processing. *Proc. Natl Acad. Sci. USA*, **103**, 12707–12712.
58. Guo,M., Xu,F., Yamada,J., Egelhofer,T., Gao,Y., Hartzog,G.A., Teng,M. and Niu,L. (2008) Core structure of the yeast spt4-spt5 complex: a conserved module for regulation of transcription elongation. *Structure*, **16**, 1649–1658.