doi:10.1016/j.jmb.2009.05.078 J. Mol. Biol. (2009) 391, 341–358

JMB

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

Two Structurally Independent Domains of *E. coli* NusG Create Regulatory Plasticity *via* Distinct Interactions with RNA Polymerase and Regulators

Rachel Anne Mooney[1], Kristian Schweimer[2], Paul Rösch[2], Max Gottesman[3] and Robert Landick[1,4]*

[1]*Department of Biochemistry, University of Wisconsin, Madison, WI 53706, USA*

[2]*Department of Biopolymers, Universitaet Bayreuth, 95440 Bayreuth, Germany*

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

[4]*Department of Bacteriology, University of Wisconsin, Madison, WI 53706, USA*

Received 6 March 2009; received in revised form 27 May 2009; accepted 29 May 2009
Available online 3 June 2009

Edited by R. Ebright

NusG is a conserved regulatory protein that interacts with elongation complexes (ECs) of RNA polymerase, DNA, and RNA to modulate transcription in multiple and sometimes opposite ways. In *Escherichia coli*, NusG suppresses pausing and increases elongation rate, enhances termination by *E. coli* ρ and phage HK022 Nun protein, and promotes antitermination by λN and in ribosomal RNA operons. We report NMR studies that suggest that *E. coli* NusG consists of two largely independent N- and C-terminal structural domains, NTD and CTD, respectively. Based on tests of the functions of the NTD and CTD and variants of NusG *in vivo* and *in vitro*, we find that NTD alone is sufficient to suppress pausing and enhance transcript elongation *in vitro*. However, neither domain alone can enhance ρ-dependent termination or support antitermination, indicating that interactions of both domains with ECs are required for these processes. We propose that the two domains of NusG mediate distinct interactions with ECs: the NTD interacts with RNA polymerase and the CTD interacts with ρ and other regulators, providing NusG with different combinations of interactions to effect different regulatory outcomes.

Keywords: NusG; RNA polymerase (RNAP); ρ termination; transcription regulation

Introduction

NusG is an exceptionally conserved regulator of gene expression that exerts complex and sometimes opposite effects on transcript elongation by RNA polymerase (RNAP). However, little is known about how NusG interacts with elongation complexes (ECs) to generate these effects on transcription. Alone, *Escherichia coli* NusG (*Ec*NusG) increases the elongation rate of *E. coli* RNAP both *in vivo* and *in vitro*,[1–3] primarily by suppressing transcriptional pauses that involve backtracking by RNAP.[3,4] However, in *Bacillus subtilis*, NusG actually enhances rather than suppresses transcriptional pausing.[5] The role of NusG in termination is also complex. *Ec*NusG enhances ρ-dependent transcription termination, whereas, in combination with NusA, NusB, NusE, and either λN or other cellular factors, NusG can modify RNAP to create specialized "antitermination" complexes that resist pausing and termination. These complexes are required for bacteriophage λ growth or for efficient transcription of ribosomal RNA operons.[6,7] This ability of NusG

**Corresponding author.* Department of Biochemistry, University of Wisconsin, Madison, WI 53706, USA. E-mail address: landick@bact.wisc.edu.

Abbreviations used: EC, elongation complex; NTD, N-terminal domain; CTD, C-terminal domain; RNAP, RNA polymerase; *Ec*NusG, *Escherichia coli* NusG; NGN, NusG N-terminal; *Tt*NusG, *Thermus thermophilus* NusG; *Aa*NusG, *Aquifex aeolicus* NusG; DSIF, DRB sensitivity-inducing factor; NELF, negative elongation factor; FL, full-length; HSQC, heteronuclear single quantum coherence; PDB, Protein Data Bank; NOESY, nuclear Overhauser enhancement spectroscopy; OD_{600}, optical density at 600 nm.

to inhibit rather than stimulate termination cannot be explained as a difference between NusG acting alone on RNAP *versus* acting as part of a multi-protein assemblage. Indeed, not all NusG-containing multi-factor complexes are resistant to termination. During λ infection of an *E. coli* HK022 lysogen, NusG, complexed with the other Nus factors and the phage HK022 Nun protein, promotes transcription termination.[8] NusG is essential in wild-type *E. coli* K12 because it enhances ρ-dependent termination upstream of the *rac* prophage gene *kil*; *nusG* can be deleted if *kil* also is deleted or inactivated.[9,10] The essentiality of NusG is species-specific: NusG is dispensable in *Staphylococcus aureus*, and both NusG and ρ are dispensable in *B. subtilis*.[11–13] This extensive and complex involvement of NusG in transcriptional processes with different transcriptional outcomes and the apparent species-specific differences in NusG activity have stymied understanding of how NusG regulates ECs. Additionally, the regions of NusG that participate in interactions with RNAP or transcription factors have not yet been determined.

X-ray or NMR structures were recently obtained for NusG from *Thermus thermophilus* (*Tt*NusG) and *Aquifex aeolicus* (*Aa*NusG),[14–17] for the NusG paralog RfaH,[18] and for the NGN (NusG N-terminal) domain of *Saccharomyces cerevisiae* Spt5.[19] All NusGs are composed of conserved N-terminal domain and C-terminal domain (NTD and CTD, respectively) connected by a flexible linker that contains an NGN motif and a KOW motif, respectively, with potential to interact with either protein or nucleic acid.[20,21] The NTD of *Tt*NusG contains a flexible loop that is replaced by a third protein domain in *Aa*NusG and is absent in Spt5. Based on analogy to *Ec*RfaH[18] and on bacterial two-hybrid interaction assays,[22] the NusG NTD is proposed to bind RNAP *via* interaction with the loop between two antiparallel helices in the β′ clamp domain (the β′ clamp helices; β′CH). A hydrophobic pocket on the NusG NTD is the proposed site of β′CH interaction.[18] However, neither this proposal nor any other function of NusG's NTD has been tested experimentally.

The CTD of NusG forms a Tudor domain in *Tt*NusG and *Aa*NusG, within which is embedded the KOW sequence motif. Based on similarities to other KOW or Tudor-domain-containing proteins, the NusG CTD is proposed to interact with either protein or nucleic acid.[15] Despite the large number of potential interaction partners to mediate the plethora of NusG activities, no interaction partners with the CTD have been identified. Interestingly, the crystal lattice of one *Aa*NusG structure generates a domain-swapped dimer formed by two NusG molecules.[16] Based on this structure and strong evidence that NusG is a monomer,[4,16] Knowlton *et al.* proposed that the NusG CTD binds to the NTD to create a spring-loaded state that is released upon interaction with RNAP.[16] Although such a two-state model appears to operate for *Ec*RfaH (which has a structurally non-homologous CTD), NMR analysis of the *Tt*NTD with and without CTD revealed no changes in NTD NMR signal and, thus, does not lend support to the existence of the intense CTD–NTD interaction predicted by the spring-loaded model.[17]

Better understanding of the structure and function of *Ec*NusG is important to understand how transcription is regulated in other organisms as well. NusG is highly conserved, and genes encoding NusG or close homologs exist in all eubacterial genomes sequenced to date, although the majority of what is known about NusG function results from biochemical studies of *Ec*NusG. The *S. cerevisiae* NusG homolog, Spt5, is essential for growth, and complexes of Spt5 and Spt4 stimulate elongation by both RNAPI and RNAPII.[23,24] In metazoans, Spt5 complexes with Spt4 to form the dimeric DRB sensitivity-inducing factor (DSIF).[25] Similar to bacterial NusG, DSIF also has complex effects on transcription: it mediates negative elongation factor (NELF)-dependent blocks to transcript elongation in promoter-proximal regions but can also mediate pause suppression once NELF is removed by P-TEFb phosphorylation of DSIF, NELF, and the RNAPII C-terminal repeat domain.[26]

Since the two domains of NusG may have distinct interaction partners (e.g., NTD with RNAP, CTD with ρ, Nus factors, or RNA), we hypothesized that expression of the NusG NTD or CTD separately or of full-length (FL) NusG containing substitutions that eliminated interactions of one domain without affecting the interactions of the other could inhibit wild-type function. In addition, we tested the phenotype of the various mutants in strains lacking NusG. To allow accurate structure–function analysis of these and other substitutions, we also determined NMR structures for the NTD and CTD of *Ec*NusG.

Table 1. Experimental restraints for structure calculation

	NusG NTD	NusG CTD
Distance restraints		
Total	860	543
Intraresidual	31	21
Sequential	189	142
Medium range	254	64
Long range	386	252
Hydrogen bonds[a]	—	32
Dihedral restraints	8	17
Restraint violations		
rms distance violation (Å)	0.003	0.003 (±0.002)
Maximum distance violation (Å)	0.11	0.16
rms dihedral violation (°)	0	0.09 (±0.01)
Maximum dihedral violation (°)	0	0.5
rmsd bond length (Å)	0.0005 (±0.00004)	0.0004 (±0.00005)
rmsd bond angle (°)	0.09 (±0.003)	0.08 (±0.004)
Atomic coordinate precision		
Backbone atoms (Å)	0.54[b]	0.42[c]
All heavy atoms (Å)	0.97[b]	1.00[c]
Ramachandran plot statistics (%)[d]		
Most favored regions	89.3	83.6
Additional allowed regions	10.0	16.3
Generously allowed regions	0.5	0.1
Disallowed regions	0.2	0.0

[a] Two distance restraints per hydrogen bond.
[b] K7–T46 and Y68–Q117.
[c] Residues L127–A181.
[d] Determined by PROCHECK.