



An α Helix to β Barrel Domain Switch Transforms the Transcription Factor RfaH into a Translation Factor

Björn M. Burmann,^{1,5} Stefan H. Knauer,¹ Anastasia Sevostyanova,² Kristian Schweimer,¹ Rachel A. Mooney,³ Robert Landick,^{3,4} Irina Artsimovitch,^{2,*} and Paul Rösch^{1,*}

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

²Department of Microbiology and The RNA Group, The Ohio State University, Columbus, OH 43210, USA

³Department of Biochemistry

⁴Department of Bacteriology

University of Wisconsin, Madison, WI 53706, USA

⁵Present address: Biozentrum, University of Basel, Klingelbergstr. 70, 4056 Basel, Switzerland

*Correspondence: artsimovitch1@osu.edu (I.A.), roesch@unibt.de (P.R.)

<http://dx.doi.org/10.1016/j.cell.2012.05.042>

SUMMARY

NusG homologs regulate transcription and coupled processes in all living organisms. The *Escherichia coli* (*E. coli*) two-domain paralogs NusG and RfaH have conformationally identical N-terminal domains (NTDs) but dramatically different carboxy-terminal domains (CTDs), a β barrel in NusG and an α hairpin in RfaH. Both NTDs interact with elongating RNA polymerase (RNAP) to reduce pausing. In NusG, NTD and CTD are completely independent, and NusG-CTD interacts with termination factor Rho or ribosomal protein S10. In contrast, RfaH-CTD makes extensive contacts with RfaH-NTD to mask an RNAP-binding site therein. Upon RfaH interaction with its DNA target, the operon polarity suppressor (*ops*) DNA, RfaH-CTD is released, allowing RfaH-NTD to bind to RNAP. Here, we show that the released RfaH-CTD completely refolds from an all- α to an all- β conformation identical to that of NusG-CTD. As a consequence, RfaH-CTD binding to S10 is enabled and translation of RfaH-controlled operons is strongly potentiated.

INTRODUCTION

Regulators of the NusG family of proteins (Spt5 in yeast and Archaea, DSIF in humans) comprise the only universally conserved family of transcription factors (Werner and Grohmann, 2011). These proteins contain two common domains, the N-terminal NGN domain and one (or several in eukaryotes) C-terminal KOW domain, which are connected by a flexible linker. The NTDs apparently play the same role in Bacteria (*E. coli*), Archaea (*Pyrococcus furiosus*; *Methanococcus jannaschii*), and Eukaryota (*Saccharomyces cerevisiae*) (Hirtreiter et al., 2010; Klein et al., 2011; Martinez-Rucobo et al., 2011;

Sevostyanova et al., 2011). They directly bind to RNAP and modify it into a processive, pause-resistant state. Recent reports suggest a conserved molecular mechanism: the NTD bridges the gap between the two largest RNAP subunits and becomes part of a processivity clamp that surrounds the nucleic acid chains in the transcription elongation complex (TEC) (Hirtreiter et al., 2010; Martinez-Rucobo et al., 2011; Sevostyanova et al., 2011). The closed clamp is thought to disfavor TEC isomerization into paused and termination states, thereby ensuring that RNAP completes the synthesis of the nascent RNA. The presence of a NusG-like factor in the last universal common ancestor (Werner and Grohmann, 2011) underscores its importance for processive transcription in all living organisms.

The CTDs of NusG-like proteins interact with diverse cellular partners and may play very diverse and opposing roles, as illustrated by the best-studied members of this family, *E. coli* NusG and RfaH. NusG is associated with RNAPs transcribing most operons (Mooney et al., 2009a) and assists a subset of Rho termination events including termination of defective mRNAs and spurious transcripts (Burns and Richardson, 1995); NusG-CTD interacts with termination factor Rho to induce premature RNA release. In contrast, RfaH is only recruited to horizontally transferred operons that contain a 12 nucleotide (nt) operon polarity suppressor (*ops*) site located in their leader regions (Belogurov et al., 2009). These operons encode biosynthesis of lipopolysaccharide (LPS) core, O-antigen, capsules, and hemolysin (Bailey et al., 1997) and are dramatically activated by RfaH, which decreases strong Rho-mediated polarity, in part by excluding NusG from RNAPs transcribing the *ops*-containing operons (Belogurov et al., 2009; Burns et al., 1998; Sevostyanova et al., 2011).

Comparison of RfaH and NusG structures readily explains their similarities and differences (Figure 1). The structurally similar NTDs bind to the RNAP β' clamp helices domain (Belogurov et al., 2007; Mooney et al., 2009b) and increase the elongation rate in vitro (Artsimovitch and Landick, 2000, 2002; Burova et al., 1995), whereas the CTDs are conformational antipodes. In the crystal structure, RfaH-CTD folds as an α hairpin

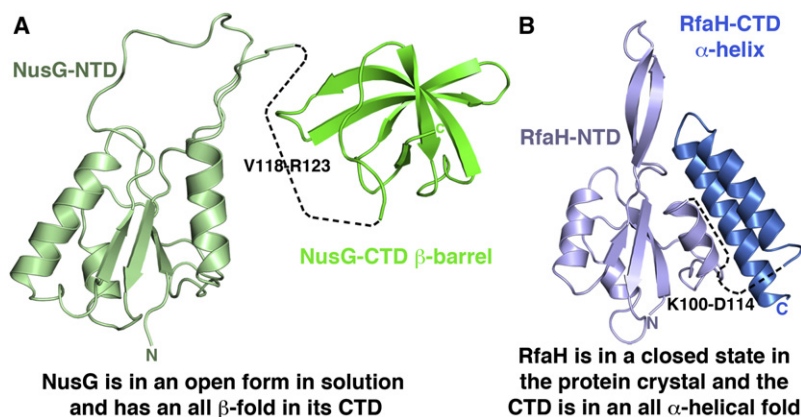


Figure 1. Structures of NusG and RfaH

(A) Ribbon representation of NusG (NTD, light green, PDB-ID: 2K06; CTD, green, PDB-ID: 2JVV; Mooney et al., 2009b).

(B) Ribbon representation of full-length RfaH (NTD, light blue; CTD, blue, PDB-ID: 2OUG (Belogurov et al., 2007), with RfaH-CTD in an α -helical fold. Termini and linker regions (broken lines) are labeled.

that tightly binds RfaH-NTD, locking the protein in a “silent” state by hiding the RNAP-binding surface of RfaH-NTD in the domain interface (Belogurov et al., 2007). When transcribing RNAP pauses at the *ops* site, interactions between RfaH-NTD, *ops* bases in the nontemplate DNA strand, and RNAP appear to trigger domain separation, thus “activating” RfaH. The isolated RfaH-NTD mediates all characteristic effects of RfaH on transcription elongation but no longer requires the *ops* site for recruitment to the TEC (Belogurov et al., 2007). Thus, the only known role of RfaH-CTD is to restrict RfaH action to *ops*-containing operons by masking its RNAP-binding surface.

NusG-CTD forms an antiparallel barrel-like β sheet (Mooney et al., 2009b) that does not stably interact with NusG-NTD in solution (Burmam et al., 2011), leaving the RNAP-binding site exposed. The residues that interact with the *ops* element in RfaH-NTD are not conserved in NusG-NTD (Belogurov et al., 2010), which is consistent with NusG association with all *E. coli* operons except those that contain *ops* (Belogurov et al., 2009). Similarly to RfaH-NTD, isolated NusG-NTD reduces RNAP pausing in vitro (Mooney et al., 2009b). However, NusG-CTD plays a much more active role. It either interacts with Rho to facilitate RNA release (Mooney et al., 2009b), binds to ribosomal protein S10 (identical to NusE) to couple transcription and translation (Burmam et al., 2010), or mediates formation of an rRNA antitermination complex (Squires and Zaporjets, 2000). A similar domain architecture is observed in NusG and Spt5 proteins from many organisms (Martinez-Rucobo et al., 2011; Mooney et al., 2009b; Reay et al., 2004; Steiner et al., 2002; Zhou et al., 2009a), and the CTDs of these proteins interact with diverse partners (Chen et al., 2009; Schneider et al., 2006; Zhou et al., 2009b), suggesting that the β barrel conformation is a preferred CTD state that facilitates linkage between transcription and other key cellular processes.

Although a conformational change of this magnitude would be unprecedented, several observations led us to propose that, upon activation, RfaH-CTD may refold into a NusG-CTD-like structure and interact with the ribosome (Belogurov et al., 2009). First, the RfaH-CTD sequence can be easily integrated into a β barrel structure (Belogurov et al., 2007). Second, RfaH-CTD capture by the trailing ribosome could explain why RfaH is stably bound to TEC in vivo, but not in vitro (Belogurov et al., 2009). Third, RfaH effects on gene expression in vivo are much

larger than those on transcription in vitro (Sevostyanova et al., 2011), suggesting the influence of additional factors. Finally, RfaH associates directly or indirectly with the translational machinery (Bailey et al., 2000). Here, we show that RfaH-CTD can indeed switch its protein fold from all- α to all- β , enabling interactions with S10 and converting RfaH into a potent activator of translation.

RESULTS

Full-Length RfaH Maintains a Closed State in Solution

To test whether the RfaH-CTD exhibits different folds in the crystal and in solution, we assigned the backbone atoms of the full-length protein in solution nuclear magnetic resonance (NMR) spectra. NMR ^1H and ^{13}C secondary chemical shifts match the secondary structure elements of RfaH found in the crystal structure (Figures 2A and 2B), and both RfaH-CTD helices are well defined, ruling out an equilibrium of different RfaH-CTD conformations in solution in the closed state. Nevertheless, helix $\alpha 4$ appears to be one turn shorter in solution than in the crystal, likely due to helix destabilization by G121. Chemical shifts are close to random coil values for K102–G121, suggesting high conformational flexibility, consistent with the lack of observable electron density in the crystal structure. Backbone amide resonances could not be detected for R11, G12, R40, R43, T66, T67, R73, G82, and N156, which are all located in loop regions in the crystal structure and may exhibit line broadening due to fast solvent exchange in solution.

To test whether RfaH-NTD:RfaH-CTD interactions are maintained in solution, we determined the rotational correlation time τ_c for full-length RfaH. ^{15}N transversal (R_2) and longitudinal (R_1) relaxation rates at 18.8 T and 288 K (Figures 2C and 2D) were $R_1 = 0.68 \pm 0.12 \text{ s}^{-1}$ and $R_2 = 31.8 \pm 5.2 \text{ s}^{-1}$ (excluding flexible regions), thus $\tau_c = 13.4 \text{ ns}$ for an isotropic model. These values are characteristic for monomeric proteins the size of RfaH. The R_1/R_2 distributions of RfaH-NTD and RfaH-CTD (Figure 2E) suggest that both domains reorient with identical correlation times, indicating tight domain interaction as in the crystal. Lower R_1/R_2 ratios reflect subnanosecond flexibility for the linker region. These results establish that the inactive RfaH conformation captured in the crystal is preserved in solution.

RfaH-CTD Refolds into a β Barrel upon Release from RfaH-NTD

In $[^1\text{H}, ^{15}\text{N}]$ heteronuclear single-quantum coherence (HSQC) spectra for isolated RfaH-CTD, we observed chemical shifts differing significantly from those of the full-length protein

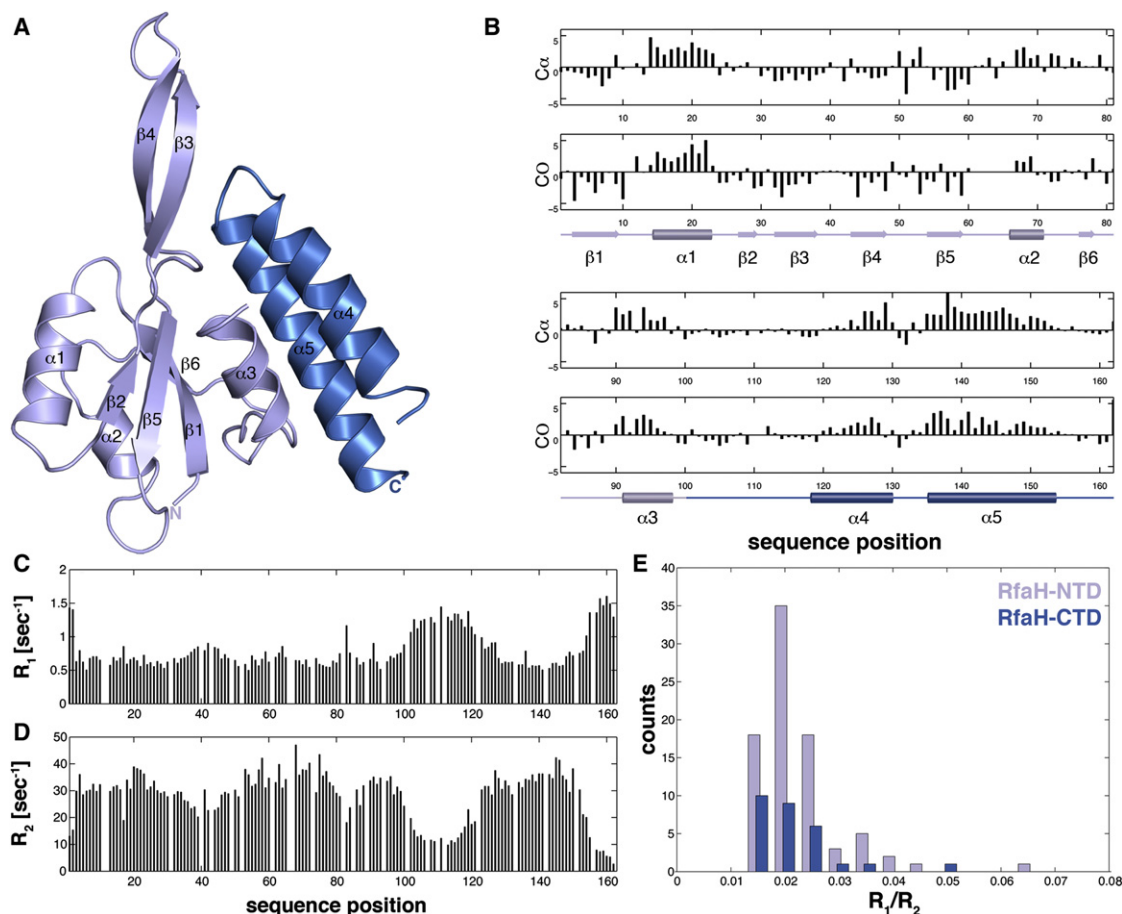


Figure 2. Structure of Full-Length RfaH in Solution

(A) Ribbon representation of the crystal structure. PDB-ID: 2OUG (Belogurov et al., 2007). NTD, light blue; CTD, blue. Termini and secondary structure elements are labeled.

(B) Chemical shift index (CSI) for $C\alpha$ and CO versus sequence position. Secondary structure elements as in (A).

(C) R_1 for full-length RfaH.

(D) R_2 for full-length RfaH.

(E) R_1/R_2 is unimodal for full-length RfaH. RfaH-CTD (124–158), blue bars; RfaH-NTD (1–100), light blue bars.

(Figure S1 available online). Consistent with an all- α to all- β rearrangement, the isolated RfaH-CTD does not contain any helical structure but consists of five β strands, K115–I118, Q127–F130, R138–N144, E149–K155, and F158–K160, which form a well-defined antiparallel β sheet with strand order $\beta 5$ – $\beta 1$ – $\beta 2$ – $\beta 3$ – $\beta 4$ (Figures 3A and 3B and Table S1). Strikingly, the folds of isolated RfaH-CTD and NusG-CTD are virtually identical in solution (backbone root mean square deviation: 0.65 Å for P112–L162). The largest local structural difference between the two CTDs is in the region corresponding to E132–R138 in RfaH, which lacks interstrand contacts typical for sheets, whereas the comparable region in NusG-CTD forms strand $\beta 2$. In RfaH, this region is 1 residue shorter and contains P133, which has no equivalent in NusG. Thus, isolated, separately expressed RfaH-CTD exists as a β barrel in solution. This result predicts that RfaH-CTD completely refolds from the α hairpin, stabilized by contacts with RfaH-NTD, when the domain interface is disrupted during RfaH recruitment to the TEC.

We next asked whether this dramatic structural rearrangement can occur within the full-length protein. Because we were unable to observe different folds of RfaH-CTD within the full-length RfaH in solution, we tested conditions that were expected to destabilize the NTD:CTD interface—specifically, elevated temperature or presence of trifluoroethanol. Both led to immediate, complete protein precipitation, most likely due to exposure of the large nonpolar interdomain surface.

We next tried weakening the RfaH-NTD:RfaH-CTD interaction while maintaining protein solubility by E48S substitution in the full-length protein. This substitution is predicted to break the salt bridge E48:R138 that connects the two domains. In the $[^1H, ^{15}N]$ -HSQC spectrum of the E48S variant, we could clearly observe signals of the RfaH-NTD and, simultaneously, α -helical and β -barrel-type RfaH-CTD (Figures 4A and 4B). These results demonstrate that the structural switch of RfaH-CTD may occur within the full-length protein with destabilized interdomain contacts. As evidenced by similar intensities of the peaks

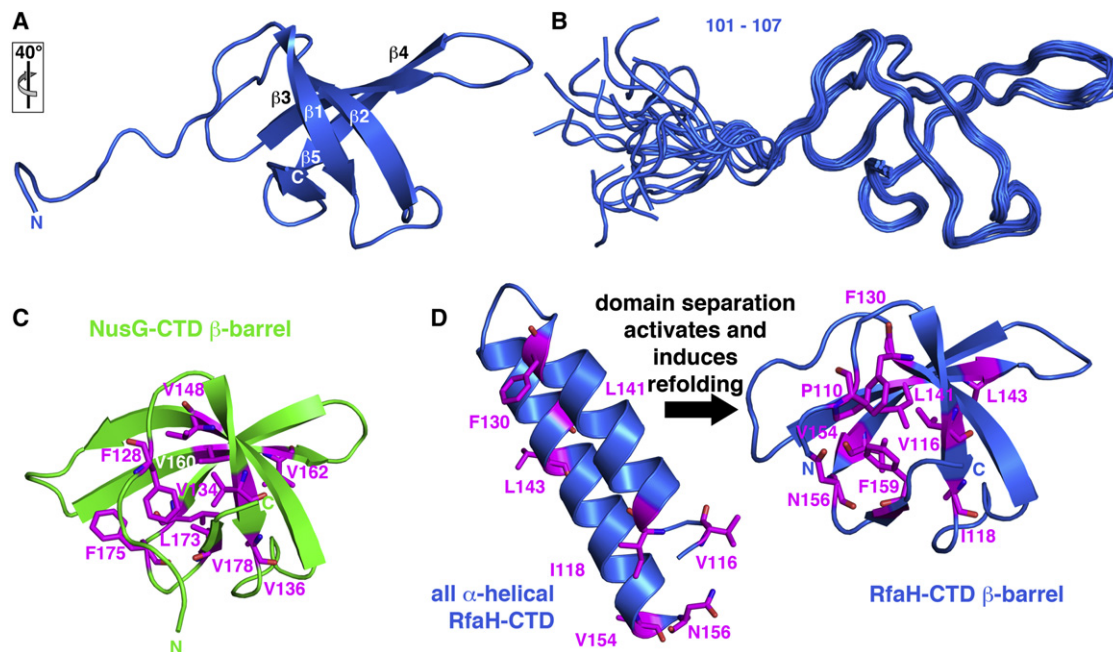


Figure 3. Structural Transition of RfaH-CTD

(A) Ribbon diagram of a representative RfaH-CTD structure, residues 101–162; termini and secondary structure elements are labeled. Orientation relative to Figure 1 is indicated.

(B) Ensemble of 20 accepted lowest energy structures. The flexible part, residues 101–107, is indicated.

(C) Ribbon representation of NusG-CTD (PDB-ID: 2JVV; Mooney et al., 2009b). Residues in the hydrophobic core of the CTD, sticks. Carbon, magenta; oxygen, red; nitrogen, blue; sulfur, yellow.

(D) RfaH-CTD (PDB-ID: 2OUG; Belogurov et al., 2007) in the all- α -helical conformation; residues corresponding to the NusG-CTD core, magenta sticks. Within the α -helical RfaH-CTD, these residues are scattered randomly along the helices (middle). For residues P110 and F159, no electron density could be determined in the full-length protein (Belogurov et al., 2007); therefore, these are excluded. Domain opening leads to drastic refolding of RfaH-CTD (right) into an all- β -barrel state; the residues corresponding to the NusG-CTD core also form the refolded RfaH-CTD core.

See also Table S1 and Figures S1 and S2.

originating from α helix and β barrel, the two conformations of RfaH-CTD in the E48S variant coexist in roughly 1:1 equilibrium.

As the E48S substitution is located in RfaH-NTD, it is not expected to exert a direct effect on the conformation of RfaH-CTD. However, to verify that RfaH-CTD can refold in a protein with unmodified domains, we used an RfaH variant with a TEV cleavage site engineered into the linker. The linker is highly flexible and tolerates deletions and insertions (I.A., unpublished data), and its sequence is not conserved (Carter et al., 2004); thus, the functional properties of RfaH(NTD-TEV-CTD) are identical to those of wild-type (WT) RfaH (Figure 4C; Belogurov et al., 2007). The linker nonetheless serves as a flexible tether that restricts uncorrelated domain diffusion, thus enhancing the collision probability of the two domains and stabilizing their interactions. To induce domain separation, we added trace amounts of TEV protease and recorded $[^1\text{H}, ^{15}\text{N}]$ -HSQC spectra. Initially, we observed a spectrum typical for full-length RfaH, except for additional signals from the seven residues of the TEV site in the linker region, as expected. After incubation for 42 hr, we were able to detect the signals of the RfaH-CTD in the β barrel conformation. Simultaneously, signals for the α -helical RfaH-CTD disappeared due to refolding, and signals for RfaH-NTD were lost due to precipitation. The observed refolding rate is

likely limited by slow TEV-mediated cleavage of the linker under our experimental conditions.

These results clearly demonstrate that RfaH-CTD refolds spontaneously when released from RfaH-NTD. Because the two domains must be separated to expose the RNAP-binding site on RfaH-NTD, our observations strongly suggest that this radical transformation occurs in the context of the WT protein during transcription.

RfaH-CTD is able to adopt two entirely different structures with completely different internal amino acid contacts. In the domain-closed form of RfaH, the large hydrophobic interdomain surface (Belogurov et al., 2007) may foster denser packing of the RfaH-CTD than would be possible in a β fold, whereas domain opening allows refolding of the RfaH-CTD into a β fold and exposure of the hydrophobic RfaH-NTD interdomain surface that also serves as an RfaH-NTD:RNAP interface (Figure S2). Interestingly, the hydrophobic RfaH-CTD interdomain residues are not present exclusively in the hydrophobic core of the β barrel (Figure S2): I129, L142, and I146 are outside of the barrel, whereas V116, V154, and F159, the key residues of the RfaH-CTD's hydrophobic core in β barrel conformation, are surface accessible or are located in flexible regions in the α -helical conformation of full-length, domain-closed RfaH (Figures 3C and 3D).

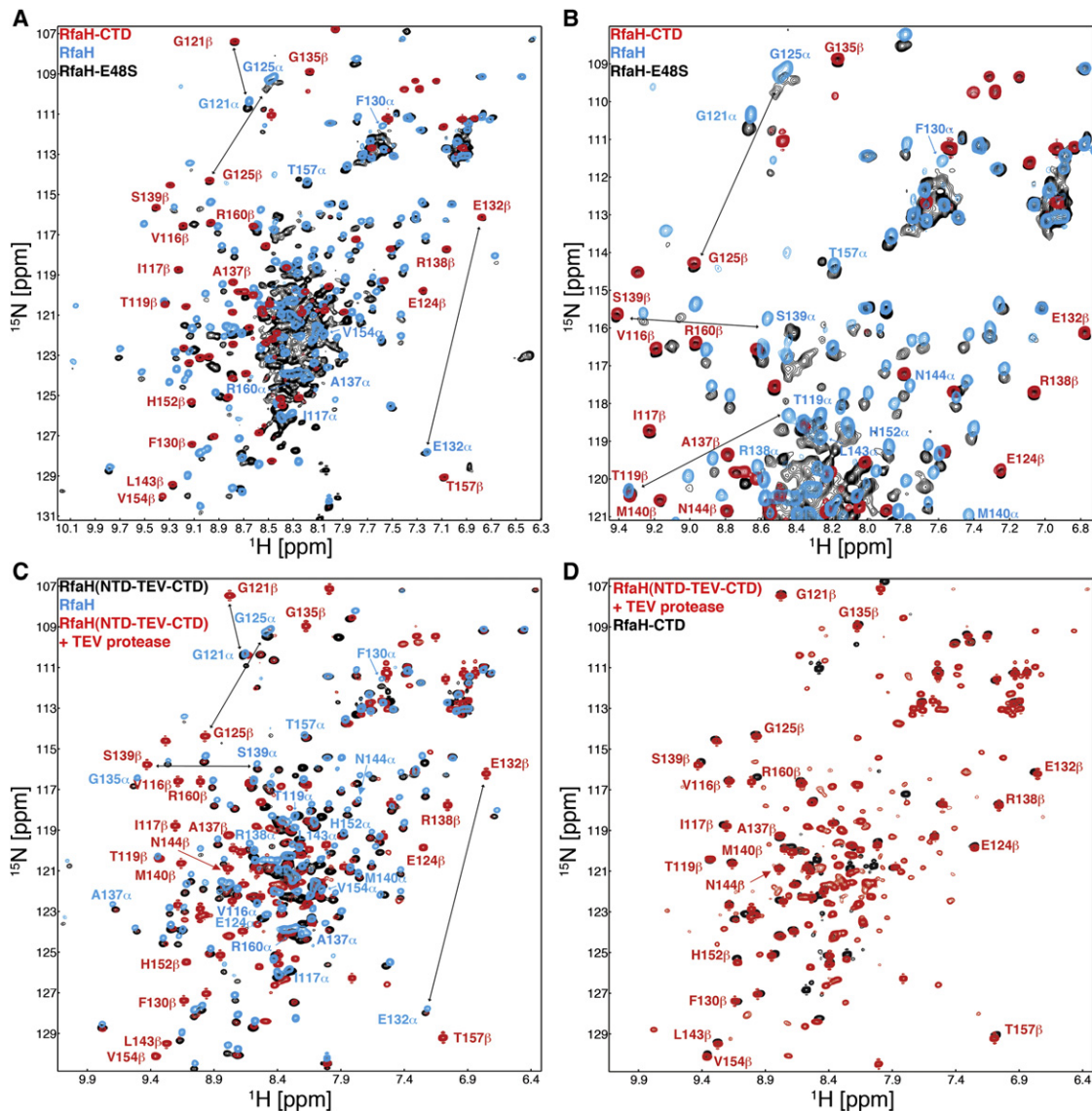


Figure 4. Structural Transition of RfaH-CTD in Full-Length RfaH

(A and B) Overlay of $[^1\text{H}, ^{15}\text{N}]$ -HSQC spectra of ^{15}N -RfaH-E48S 45 μM , black; ^{15}N -RfaH, 189 μM , cyan; and ^{15}N -RfaH-CTD, 344 μM , red. Full spectra (A) and enlargement of one region (B). Signals from the CTD in full-length RfaH (α -helical form), cyan; from the isolated CTD (β barrel form), red.

(C) Overlay of $[^1\text{H}, ^{15}\text{N}]$ -HSQC spectra of ^{15}N -RfaH(NTD-TEV-CTD), 138 μM , black; ^{15}N -RfaH, 189 μM , cyan; and ^{15}N -RfaH(NTD-TEV-CTD) after incubation with 1.75 μM TEV protease for 42 hr, 127 μM , red.

(D) Overlay of $[^1\text{H}, ^{15}\text{N}]$ -HSQC spectra of ^{15}N -RfaH(NTD-TEV-CTD) 127 μM , red. After incubation with 1.75 μM TEV protease for 42 hr from (C) and ^{15}N -RfaH-CTD, 344 μM , black; signals of the isolated CTD in the β barrel conformation, red.

See also Figures S3 and S4.

Functional Role for RfaH-CTD Refolding

Prior to engagement with transcribing RNAP, the α hairpin RfaH-CTD interacts with RfaH-NTD to restrict RfaH action to *ops*-containing operons, avoiding interference with gene regulation by housekeeping NusG (Belogurov et al., 2009). In NusG, as well as in its archaeal and eukaryotic homologs, the β barrel establishes functional interactions with diverse cellular targets. To address the impact of RfaH-CTD refolding, we wanted to determine whether it merely enables formation of a more stable

β barrel or plays a further role in the control of gene expression by RfaH.

Expression of the distal genes in an RfaH-dependent *rfb* operon is nearly abolished by Rho in a $\Delta rfaH$ strain, suggesting that the *rfb* operon is poorly translated; indeed, it displays a very high fraction of rare codons and lacks a strong translation-initiation signal (Figure S3). RfaH restores expression of distal genes, e.g., *wbbI* (Sevostyanova et al., 2011), but does not prevent Rho association throughout the *rfb* operon

(Belogurov et al., 2009). RfaH could inhibit Rho activity by: (1) modifying the TEC properties; (2) excluding NusG from RNAP, thereby inhibiting Rho-mediated RNA release (but not Rho recruitment, which is unaffected by NusG); (3) binding to Rho in an unproductive fashion; or (4) increasing translation. Our previous data are consistent with the first two mechanisms (Belogurov et al., 2009; Sevostyanova et al., 2011).

We next asked whether β barrel RfaH-CTD could interact with Rho (Figure S4) similarly to NusG-CTD (Mooney et al., 2009b). The $[^1\text{H}, ^{15}\text{N}]$ -HSQC spectrum of ^{15}N -labeled RfaH-CTD showed no significant changes on titration with unlabeled Rho, indicating that RfaH-CTD does not bind to Rho under conditions in which NusG-CTD forms a Rho complex. As no structural data for the NusG-CTD:Rho complex are available, a detailed explanation of this difference is impossible. Mutagenesis and biochemical data, however, indicate that the loops between strands $\beta 1/\beta 2$ and $\beta 3/\beta 4$ are essential for NusG:Rho interactions because deletions in these regions abrogate complex formation as assayed by in vitro pull-down experiments (Chalissery et al., 2011). Several point mutations in these regions affect Rho termination efficiency; among these are L158Q, S163F, and G166D in NusG (Chalissery et al., 2011; Mooney et al., 2009b; Sullivan et al., 1992), which correspond to S139, N144, and N147, respectively, in RfaH. The different characteristics of these residues may explain the inability of RfaH to interact with Rho.

The dramatic RfaH effect on Rho-mediated polarity in vivo (Sevostyanova et al., 2011), as compared with a modest effect in vitro (Belogurov et al., 2009), and the lack of strong ribosome-binding sites (RBS) preceding the first open reading frame (ORF) in RfaH-regulated operons led us to ask whether RfaH recruits the ribosome in the absence of a strong RBS via interaction between S10 and β barrel RfaH-CTD as with NusG (Burmam et al., 2010), facilitating transcription-translation coupling.

Efficient translation may confer sufficient protection against Rho to obviate any role of RfaH in transcription-translation coupling; conversely, inhibiting translation would confer dependence on RfaH. To test these ideas, we used reporters in which a *lux* operon is positioned downstream from an *ops* element. In the presence of a strong RBS in front of the first *luxC* gene, RfaH increases luciferase signal more than 6-fold (Belogurov et al., 2010). Removal of RfaH-CTD further increased expression ~ 1.4 -fold (Figure 5A), likely because isolated NTD can be recruited to RNAP at multiple sites, in contrast to WT RfaH that becomes engaged only at an *ops* element. Alterations that compromise RfaH binding to the TEC decrease luciferase activity of this reporter (Belogurov et al., 2010); this pattern is consistent with principal effects on transcription.

To test whether compromised translation would increase dependence on RfaH, we removed the *luxC* RBS. This change eliminated *lux* expression in the absence of RfaH (Figure 5B). However, WT RfaH expressed from a plasmid increased luciferase activity 1,000-fold, restoring it to 18% of that observed with an RBS. This suggests that RfaH can indeed connect ribosomes to TECs and thereby couple translation to transcription. If this interpretation were correct, stimulation of luciferase activity should be dependent on RfaH-CTD. Indeed, the RfaH-CTD

deletion reduced *lux* expression in the absence of RBS more than 6-fold (Figure 5B).

The stimulatory effect of RfaH-NTD alone was also more pronounced in the absence of RBS. This difference is likely due to a greater dependence on the anti-Rho activity of NTD when translation is compromised and Rho can access the nascent RNA. In the presence of the RBS, Rho action will be inhibited by translation, reducing the requirement for RfaH-NTD.

These results suggest that destabilization of the interdomain interface in full-length RfaH will augment its effect on gene expression because the effects of RfaH-NTD and RfaH-CTD will be combined. To test this prediction, we constructed an E48A substitution that would be expected to disrupt the interdomain contacts, but not RfaH binding to RNAP. Indeed, we found that E48A protein was highly active in reducing RNAP pausing in vitro in the absence of an *ops* site, mimicking the effect of the RfaH-CTD deletion (Figure S5); consistently, a substitution for Ser induced RfaH-CTD refolding (Figure 4). In contrast to the isolated RfaH-NTD, the E48A variant increased reporter activity in vivo (Figures 5A and 5B); the E48A variant was ~ 4 -fold more active than the WT RfaH in the absence of RBS. These results are consistent with the direct role of the refolded RfaH-CTD in stimulation of *lux* operon translation.

NusG:S10 interactions are mediated by hydrophobic residues (Burmam et al., 2010), some of which are conserved in the RfaH-CTD (Figures S6A–S6C). NusG F165 (I146 in RfaH) is critical for this interaction, is highly conserved (Burmam et al., 2010), is required for viability (Knowlton et al., 2003), and is crucial for NusG:S10 interactions in vivo (Mooney et al., 2009b). RfaH I146D substitution (to eliminate putative van der Waals contacts with S10) strongly reduced *lux* expression in the absence, but not in the presence, of RBS (Figures 5A and 5B), supporting a model in which RfaH-CTD:S10 interactions are essential for translation activation.

We next tested whether this pattern holds for the RfaH-controlled *rfb* operon. Consistent with poor translation, expression of the eighth gene (*wbbI*) lacking an RBS is nearly abolished by Rho in the absence, but not presence, of RfaH (Sevostyanova et al., 2011). The effects of RfaH variants on *wbbI* expression were similar to those observed with the RBS-deficient *lux* operon (Figure 5C). The E48A substitution increased *wbbI* mRNA levels relative to WT protein 1.6-fold, whereas RfaH-CTD deletion and I146D substitution decreased them ~ 11 - and 7-fold, respectively. L145D, a substitution of another residue at the putative RfaH-CTD:S10 interface, had a less dramatic but still significant effect (~ 4 -fold). Together, these data argue that RfaH-CTD mediates efficient expression of translationally challenged operons that carry *ops* elements. Two residues that interact with S10 in NusG-CTD (Burmam et al., 2010) are critical for this effect, strongly indicating that the β barrel form of RfaH-CTD establishes similar contacts with S10.

RfaH Interacts with S10 In Vivo

To confirm that RfaH mediates S10 interaction in vivo, we used targeted ChIP-chip to compare the relative levels of S10 (NusE), RNAP, RfaH, NusG, and NusB on the *rfb* operon (Figure 5D) to those on a conventional protein-coding operon, *atp* (Figure 5E), an rRNA operon (Figure 5F), and two sRNA genes,

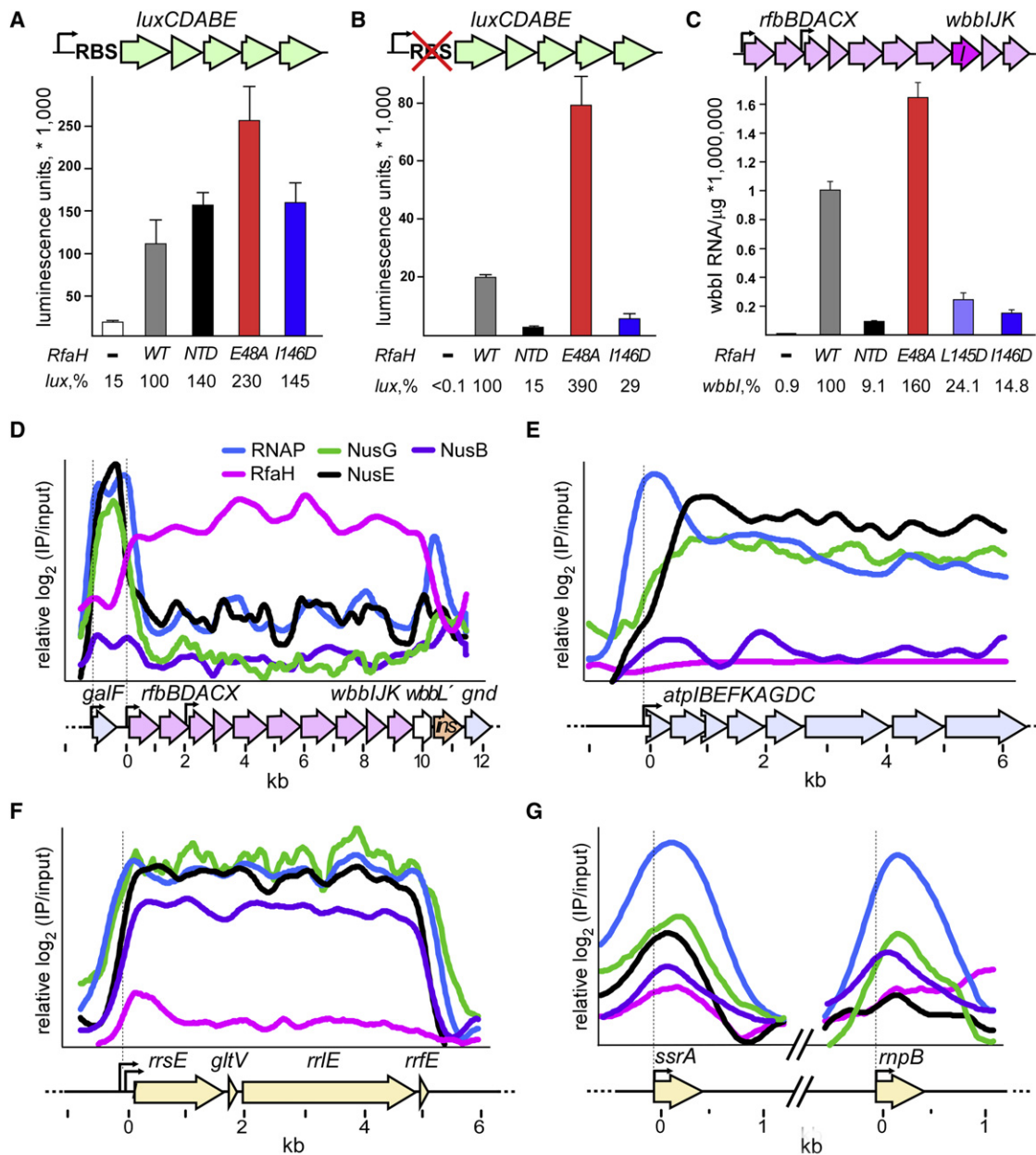


Figure 5. RfaH-CTD Is Required for Effects on Translation and May Interact with S10

(A and B) Reporter assays using the translation-competent (A) (pIA955) and defective (B) (pIA1087) *ops-lux* operon constructs, which differ in sequence 5 nt upstream from the ATG codon, GAGGA and CACAC, respectively. The assays were performed in the *rfaH* deletion strain transformed with plasmids encoding RfaH variants under control of the P_{BAD} promoter, as before (Belogurov et al., 2010). The results are expressed as luminescence corrected for the cell densities of individual cultures.

(C) Expression of *rfaH* operon (top) was evaluated by qRT-PCR. Total RNA was isolated from $\Delta rfaH$ cells expressing WT or an altered RfaH variant, and the absolute amount of *wbbI* mRNA was measured. In (A)–(C), errors ($\pm$ SD) were calculated from three independent experiments.

(D–G) ChIP-chip analysis of the protein-coding *rfaH* and *atp* operons and noncoding *rrsE*, *mpB*, and *ssrA* genes, performed as described previously (Belogurov et al., 2009; Mooney et al., 2009a) with probe sets for the *rfaH* (D), *atp* (E), *rrsE* (F), *ssrA* (G), and *mpB* (G) transcription units (TUs). Cy3 signal (IP) from the DNA immunoprecipitated with monoclonal antibodies to RNAP (β subunit), NusA, or HA epitope tag on NusG or polyclonal antibodies to RfaH, NusE (S10), or NusB was divided by Cy5 signal (input) from unenriched DNA collected prior to immunoprecipitation. The data for each target were quantile normalized against each other so that relative signal ratios could be compared and plotted on a log₂ scale. The ratios of average NusE/RNAP and NusG/RNAP signals were 1.1 and 0.47 (*rfaH*), 1.1 and 0.88 (*atp*), 0.98 and 1.1 (*rrsE*), 0.35 and 0.45 (*mpB*), and 0.36 and 0.55 (*ssrA*).

See also Figure S5.

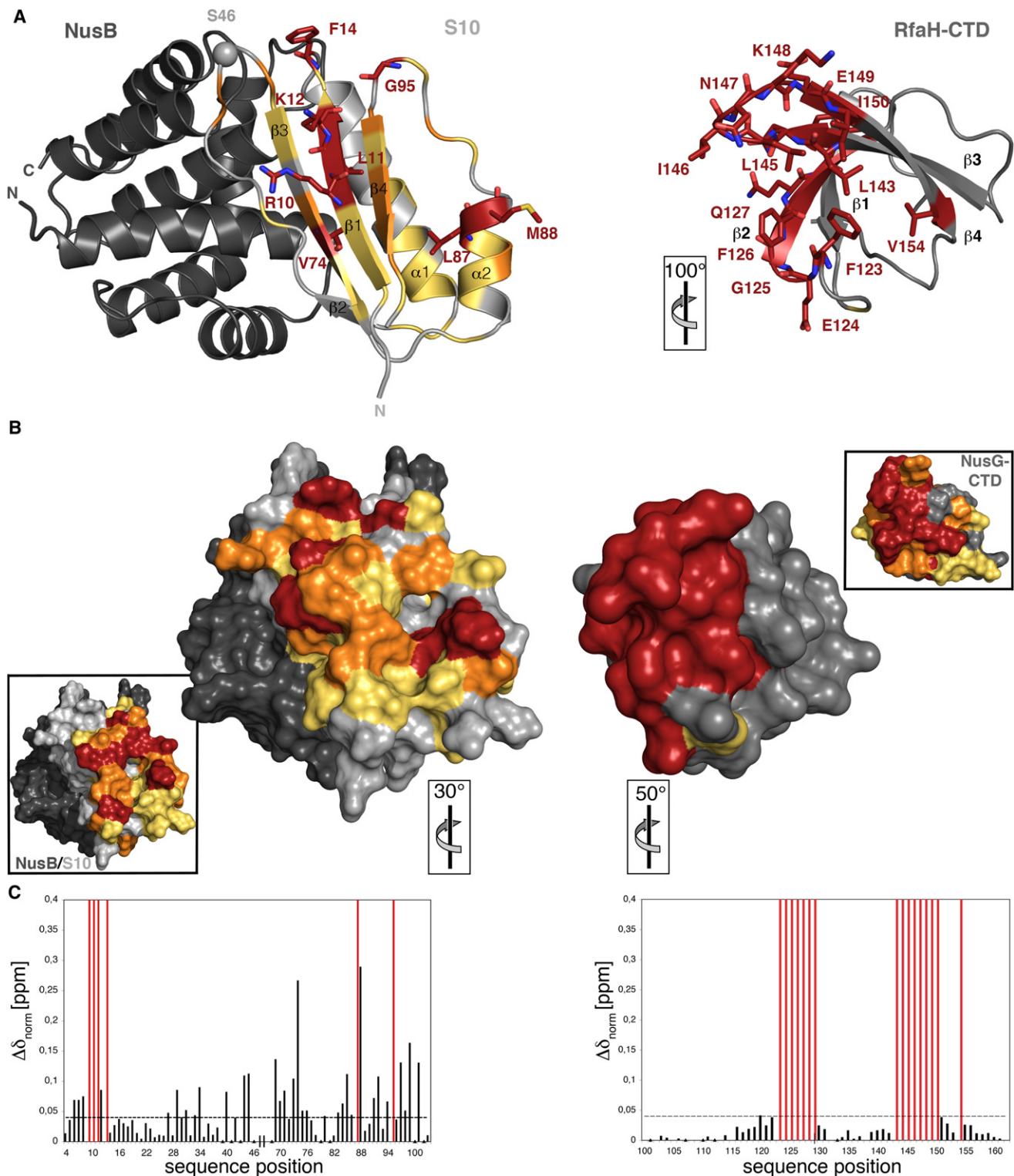


Figure 6. RfaH-CTD:S10 Interface

(A) Mapping of the titration induced ^1H , ^{15}N -HSQC chemical shift changes ($\Delta\delta_{\text{norm}}$ [ppm] > 0.2, red; > 0.1, orange; and > 0.04, yellow) on structures of the NusB:S10 complex (dark and light gray, respectively; PDB-ID 3D3B (Luo et al., 2008)) and RfaH-CTD (gray). Strongly affected residues (sticks; carbon, red; oxygen, red; nitrogen, blue; sulfur, yellow) are shown. Gray sphere in S10 denotes the C α position of S46, which in this construct replaces residues 46–67 of the WT S10 (Luo et al., 2008).

rnpB and *ssrA* (Figure 5G). Relative to RNAP signal, *rfb*, *atp*, and *rnnE* all exhibited high levels of S10, whereas the untranslated sRNA genes exhibited much lower levels of S10. The S10 signal on *rfb* likely reflects ribosomes interacting with RfaH because both NusG and NusB, the known partner of S10 in antitermination complexes (Squires and Zaporozets, 2000), were absent. In contrast, *rnnE*, on which NusB:S10-containing antitermination complexes are known to form, gave significant NusB signal. Because NusG is specifically excluded from the *rfb* operon by RfaH (Belogurov et al., 2009), these data suggest that S10 may instead interact with RfaH.

To detect the RfaH:S10 interaction in cellular extracts, we analyzed proteins that associate with RfaH-CTD using formaldehyde crosslinking followed by mass spectrometry. 30S ribosomal proteins and S10 in particular were preferentially bound to RfaH-CTD, along with TraT, an outer-membrane lipoprotein encoded by the RfaH-controlled *tra* operon (Figure S7).

RfaH:CTD Forms a Complex with S10 In Vitro

To distinguish between direct and indirect RfaH-CTD:S10 interactions, we turned to in vitro analysis. Direct RfaH-CTD:S10 contact was observed by NMR and implied by gel filtration (Figures S6D–S6F). Determination of the K_d for the RfaH-CTD:S10:NusB complex was not possible with NMR data (K_d for NusG-CTD:S10:NusB: 50 μ M; Burmann et al., 2010). Chemical shift mapping of RfaH-CTD:S10:NusB interactions suggests that RfaH-CTD and NusG-CTD have similar binding modes (Figure 6). The interaction surface is composed of a hydrophobic pocket of helix α 2 and strands β 1 and β 4 of S10 and the loops between strands β 1/ β 2, β 3/ β 4, and residues from strand β 4 of RfaH-CTD. RfaH-CTD thus forms a plug that fits into an S10 pocket, in analogy to the NusG:S10 complex (Burmann et al., 2010). The high similarity between these two interaction surfaces, together with sequence homology in the S10-binding regions of RfaH and NusG (Figure S6), indicate similar functional contacts between the β barrel CTDs and S10 and suggest that RfaH-CTD can bind to free as well as ribosome-bound S10 because its binding site remains accessible when it is part of the ribosome. Although we could only detect RfaH-CTD:S10 interactions at high concentrations in vitro, our observations that substitutions in RfaH located at the interface confer defects in vivo (Figure 5C) support a functional role of this contact in the cell and suggest that stable interaction occurs in the context of an RfaH-modified TEC and the ribosome.

To address the possibility that the interaction with S10 might trigger domain opening in full-length RfaH, we titrated unlabeled RfaH into ^{15} N-labeled S10 (Figure S6G). Interaction of the two proteins could not be observed, clearly demonstrating that S10 recognizes RfaH-CTD in the β barrel fold only and that presence

of S10 alone does not induce RfaH domain release or RfaH-CTD refolding.

DISCUSSION

RfaH strongly inhibits Rho-dependent termination via antipausal modification of RNAP and exclusion of the Rho cofactor NusG (Sevostyanova et al., 2011). Here, we present evidence that RfaH also binds to the ribosome to dramatically activate translation. This activity requires complete refolding of RfaH-CTD from α helix into β barrel conformation. We show that RfaH-CTD, which folds as an α helix in the full-length, domain-closed form of the protein, folds as a β barrel when expressed separately. Even more strikingly, in full-length RfaH, RfaH-CTD refolds to a β barrel when its interaction with RfaH-NTD is weakened by an E48S substitution or when the domain linker is broken. In vivo, the domain separation is induced by RfaH interactions with the *ops* DNA element presented in the context of the TEC wherein RfaH-NTD exchanges its contacts with the CTD for those with the β' clamp helices domain of RNAP. The dramatic conformational switch of RfaH-CTD enables RNAP coupling to a ribosome by RfaH, thus activating translation and blocking Rho-dependent termination in RfaH-controlled operons (Figure 7).

RfaH as a Transcription Antiterminator

RfaH was initially described as a transcription factor that targets a group of *ops*-containing genes in enteric bacteria. Although some of these genes encode important commensal functions (e.g., LPS core biosynthesis), most RfaH-controlled genes are essential for virulence (e.g., α -hemolysin and capsule biosynthesis). RfaH dramatically increases expression of distal genes in these operons by supporting RNAP readthrough termination signals. RfaH has thus been termed an antiterminator.

Antitermination modification of RNAP is particularly important in eukaryotes, in which transcribed units can consist of many nucleotides, and in foreign (e.g., phage) bacterial genes, which are poorly translated and subject to polarity control by joint action of Rho and NusG. All genes that are controlled by RfaH have been acquired through horizontal transfer and are thus highly sensitive to NusG-enhanced Rho-dependent termination. Consistently, RfaH's main activity is to inhibit Rho (Belogurov et al., 2009), and here it has to work against NusG, which is present in *E. coli* in concentrations >100-fold higher than RfaH is (Li et al., 1993; Belogurov et al., 2009).

Antitermination was thought to explain all transcription activation effects by RfaH. We showed that RfaH antagonizes Rho activity by two mechanisms, both mediated by the RfaH-NTD. First, RfaH competes with NusG for binding to the β' clamp helices and excludes NusG from the TEC both in vivo and in vitro, thereby reducing Rho efficiency. Second, RfaH inhibits

(B) Surface representation of the structures shown in (A). Orientations relative to (A) (NusB:S10) and Figure 1 (RfaH-CTD) are indicated. Inserts show the corresponding interaction surface of the NusB:S10:NusG-CTD complex (Burmann et al., 2010).

(C) $[^1\text{H}, ^{15}\text{N}]$ -HSQC-derived chemical shift changes versus sequence position. (Left) S10 chemical shift changes on titration with RfaH-CTD; missing residues of the S10 ribosome-binding loop are indicated by a break on the sequence axis. (Right) RfaH-CTD chemical shift changes on titration with S10. Dotted line, significance level of $\Delta\delta_{\text{norm}}[\text{ppm}] = 0.04$; red bars, signals disappearing upon complex formation. Triangles, unassigned residues.

See also Figures S6 and S7.

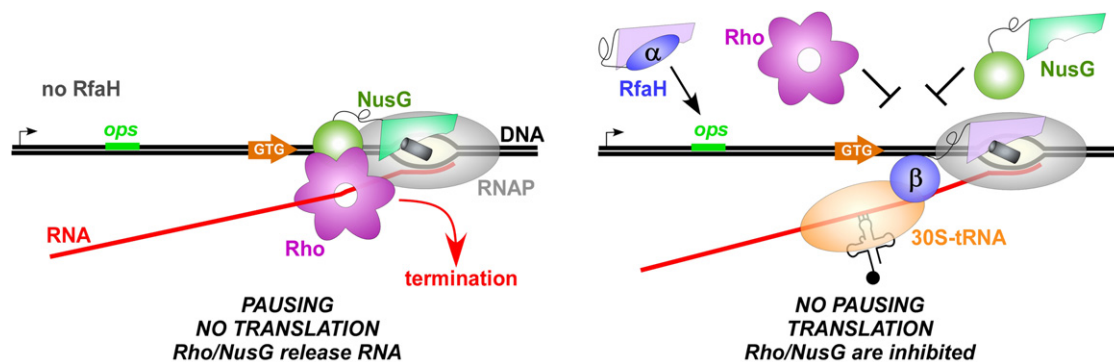


Figure 7. Model for Multifaceted Activation of Gene Expression by RfaH

In RfaH-controlled operons, the *ops* element is located within 100 nt upstream of a GTG predicted (based on protein sequence analysis) to serve as a translation start codon. In absence of RfaH (left), NusG-NTD binds to the β' clamp helices (dark gray cylinder), and NusG-CTD interacts with Rho (purple) to terminate transcription by RNAP (gray). In the *rfb* operon, Rho decreases expression of distal genes by ~ 800 -fold (Sevostyanova et al., 2011). When present (right), RfaH binds to elongating RNAP at the *ops* site and reduces Rho effect to 2-fold. This strong antipolar activity depends on the coordinated action of both RfaH domains becoming separated during recruitment. Unaltered RfaH-NTD binds to the β' clamp helices to reduce transcriptional pausing and exclude NusG-NTD from binding to RNAP; both activities inhibit Rho-dependent termination. The refolded RfaH-CTD recruits the 30S subunit (bound to the initiator tRNA) via direct contacts with S10. Thus, the tethered translation initiation complex scans the mRNA lacking a strong Shine-Dalgarno (SD) element for another yet unknown start signal. Recruitment of the ribosome directly increases translation and indirectly decreases Rho-dependent termination by shielding mRNA from Rho.

formation of the paused state, which is a target for Rho. The first mechanism is unique for RfaH, whereas the second antipassing mechanism is conserved from bacteria to humans among NusG-like proteins that share the binding site on the TEC and the molecular mechanism of antipassing modification (Hirtreiter et al., 2010; Martinez-Rucobo et al., 2011; Sevostyanova et al., 2011).

Although a combination of these two mechanisms could explain large effects of RfaH on the expression of some genes, we observe much smaller effects on transcription *in vitro* (3- to 4-fold) or on RBS⁺ *lux* reporter activity *in vivo* (7- to 10-fold). Our current study provides a solution to this puzzle: we show that RfaH dramatically stimulates expression of the *lux* reporter when its translation is compromised and that RfaH-CTD is required for this effect. In retrospect, control at the level of translation could be expected in the case of RfaH-regulated operons that have many features suggesting poor translation, from a weak translation start context to a very high fraction of rare codons (Figure S3).

RfaH as a Translation Factor

Deletion of RBS is sufficient to make the expression of a heterologous *lux* operon strongly dependent on RfaH (Figure 5). This suggests that RfaH can facilitate active recruitment of the ribosome to nascent mRNA as it emerges from an RfaH-bound RNAP. RfaH remains bound to the TEC until it reaches the operon end. Thus, if RfaH maintains contacts with the ribosome after recruitment, it will ensure tight transcription-translation coupling throughout the entire operon, thereby blocking Rho-dependent termination and possibly inhibiting ribosome pausing and release. Analysis of RfaH-dependent operons is consistent with this model. For example, the first gene in the *rfb* operon (Figure S3) lacks a canonical SD element and starts with a sub-optimal GUG, whereas several downstream genes are translationally coupled, suggesting that, once the first ribosome loads

on the mRNA, it could complete translation without dissociating from the message.

How is 30S recruitment mediated by RfaH in the absence of a match between the mRNA and 16S rRNA? We show that RfaH directly interacts with S10 and that substitutions at the RfaH:S10 interface compromise activation of translation-deficient *lux* operon and RfaH-dependent *rfb* operon. Structural modeling shows that this interaction may occur in the context of either 30S or 70S assemblies, consistent with both initial recruitment and retention of the ribosome. RfaH may induce a conformational change that allows 30S recruitment or may simply increase the local concentration of 30S in the vicinity of mRNA.

The distance between *ops* element and presumed start codon of the first gene is similar (~ 100 nt) among RfaH targets, but it is yet unknown how the ribosome-loading site is determined or whether translation starts at a unique site. We do not yet know whether RfaH (and NusG, which is typically recruited to RNAP further downstream; Belogurov et al., 2009) could also facilitate ribosome binding to suboptimal sites in the middle of an operon.

RfaH-CTD as a Metamorphous Domain

We here suggest that transcription-translation coupling may be the major mechanism by which RfaH activates gene expression, as compared to its ancient and universally conserved ability to reduce RNAP pausing. This coupling requires a dramatic rearrangement of RfaH-CTD, and RfaH-CTD stands out from a group of metamorphic proteins (Andreeva and Murzin, 2006; Tokuriki and Tawfik, 2009) as it undergoes an unprecedented complete switch from an all- α to an all- β fold, extending the classic view (Anfinsen, 1973) that a given amino acid sequence under defined environmental conditions leads to a unique three-dimensional (3D) structure. Indeed, comparison of sequentially nearly identical, rationally designed proteins reveals that only a few amino acid residues encode the secondary structure (Alexander et al., 2009; Bryan and Orban, 2010).

Structural interconversion between α helix and β sheet was mainly observed with artificial peptides by varying the solution conditions (Minor and Kim, 1996) or in myoglobin, where it was possible to induce a fold consisting of β strand fibrils into all- α -helical myoglobin under nonnative conditions (Fändrich et al., 2001). In neurodegenerative diseases like Alzheimer's and the prion-related diseases, structural rearrangement leads to formation of amyloid aggregates that consist nearly exclusively of β strands (Greenwald and Riek, 2010; Surewicz and Apostol, 2011).

Very few refolding events that lead to alternate functions, and none as dramatic as for RfaH, have been described. Lymphotactin (Ltn) exists in two active conformations under physiological conditions—as a monomer with a chemokine fold and as a dimer with a β sandwich fold—and the interconversion is induced by reorganization of an N-terminal loop to form a β strand interaction while the C-terminal α helix becomes unstructured (Tuinstra et al., 2008). The mitotic arrest deficiency 2 protein (Mad2) also exists in two conformations that are $\sim 70\%$ identical; C-terminal and N-terminal regions form an additional β hairpin as a key interaction partner binding site in a closed state, stabilizing it compared to the open, noninteracting form (Luo et al., 2004). The chloride intracellular channel 1 protein (CLIC1) is able to adopt two distinct conformations in solution, depending on its oxidation state: the formation of a disulfide-bridge mediates the conversion of the N-terminal $\alpha/\beta/\alpha$ sandwich into a three α helix bundle (Littler et al., 2010).

Although these proteins undergo significant conformational changes, the dramatic all- α to all- β switch observed with RfaH-CTD is so far unique in its magnitude, as well as in its functional consequences. Importantly, both states of RfaH-CTD play essential roles in regulation of gene expression. As an α -helical hairpin, RfaH-CTD masks the RNAP-binding surface until RfaH recognizes the *ops* site, of which fewer than 20 exist in the *E. coli* genome; this autoinhibition is essential to avoid interference with NusG. RfaH domain interaction is disrupted on RfaH-NTD recruitment to the TEC, thus enabling RfaH-CTD structural transformation. In the β barrel conformation, RfaH-CTD establishes contacts with S10, favoring ribosome recruitment. This plasticity of RfaH-CTD raises the question of whether or not the diverse interactions reported for the NusG-like CTDs of eukaryotic NusG and RfaH homologs (Chen et al., 2009; Schneider et al., 2006; Zhou et al., 2009b) are all mediated by a β fold (Klein et al., 2011; Werner and Grohmann, 2011). This view is based on the fact that all NusG-like CTDs contain a KOW motif (Kyrpides et al., 1996) that is embedded in an all- β -fold as in NusG or SPT5 proteins (Knowlton et al., 2003; Mooney et al., 2009b; Reay et al., 2004; Steiner et al., 2002), ribosomal proteins L24 (Ban et al., 2000), eL26, and eL27 (Ben-Shem et al., 2010), and the tudor domain of human SMN protein (Selenko et al., 2001). Some of these interactions may deserve closer inspection with respect to the possibility of a native α -to- β fold transition. Indeed, the ability of a protein to adopt two antagonistic folds with diverse roles is a very elegant solution to employ a single amino acid sequence for different uses. It would be surprising if this approach would not be applied by other regulatory proteins in the NusG family or other protein families to expand their repertoire of functions and interaction partners.

EXPERIMENTAL PROCEDURES

Expression and purification of proteins, experimental details of LC/MS/MS measurements, transcription pause assays, NMR measurements, and structure determination protocols are described in the [Extended Experimental Procedures](#).

Quantitative Real-Time PCR

To test the effect of mutations in RfaH on its natural target, we analyzed the expression level of *wbbI*, the eighth gene in *rfa* operon, by qRT-PCR. Vectors encoding different RfaH variants under control of P_{trc} promoter were transformed into IA149 ($\Delta rfaH$) strain (Belogurov et al., 2010) grown with agitation at 37°C to $\text{OD}_{600} = 0.6$, and protein expression was induced by addition of 0.2 mM IPTG for 2 hr. Total RNA was isolated by phenol extraction and treated with DNaseI from Epicenter (Madison, WI, USA) according to manufacturer's instructions. Control PCR with specific oligos without RT-step was performed to ensure the absence of DNA contamination. Total RNA samples (1 μg) were added to the one-step qRT-PCR reaction mix from QIAGEN (Valencia, CA, USA) and analyzed in triplets on CFX96 System from Biorad (Hercules, CA, USA). For each sample, at least three repeats in two independent experiments (starting from cell growth and RNA isolation) were performed as described (Sevostyanova et al., 2011).

In Vivo Reporter Assay

Plasmids carrying RfaH variants were cotransformed with a reporter vector carrying *lux* operon with or without RBS (pIA955 or pIA1087, respectively) into IA149 strain and plated on selective media (100 $\mu\text{g}/\text{ml}$ carbenicillin, 50 $\mu\text{g}/\text{ml}$ chloramphenicol). Single colonies were inoculated into 3 ml of LB media supplemented with antibiotics and incubated at 37°C . After 6 hr of growth, cultures were diluted into fresh LB containing antibiotics and 0.5% glucose to $\text{OD}_{600} \sim 0.01$ and were allowed to grow for an additional 2 hr. Expression of RfaH variants was induced by addition of IPTG to 0.2 mM for 3 hr. Luminescence was measured in 200 μl aliquots in triplicates on FLUOstar OPTIMA plate reader (BMG LABTECH GmbH, Offenburg, Germany) and normalized by cell density.

ACCESSION NUMBERS

Atomic coordinates for the RfaH-CTD structure were deposited with the Protein Data Bank under accession code 2LCL and RCSB ID code 102231.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Experimental Procedures, seven figures, and one table and can be found with this article online at <http://dx.doi.org/10.1016/j.cell.2012.05.042>.

ACKNOWLEDGMENTS

We thank R. Heissmann for excellent technical assistance and H.K. Seoh and C. Squires for the generous gift of antibodies used for ChIP-chip experiments. This project was supported by DFG (Ro617/18-1 to P.R.) and NIH (GM67153 to I.A.; GM38660 to R.L.).

Received: November 21, 2011

Revised: March 28, 2012

Accepted: May 7, 2012

Published: July 19, 2012

REFERENCES

- Alexander, P.A., He, Y., Chen, Y., Orban, J., and Bryan, P.N. (2009). A minimal sequence code for switching protein structure and function. *Proc. Natl. Acad. Sci. USA* 106, 21149–21154.
- Andreeva, A., and Murzin, A.G. (2006). Evolution of protein fold in the presence of functional constraints. *Curr. Opin. Struct. Biol.* 16, 399–408.

- Anfinsen, C.B. (1973). Principles that govern the folding of protein chains. *Science* 181, 223–230.
- 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.
- 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.
- Bailey, M.J.A., Hughes, C., and Koronakis, V. (1997). RfaH and the *ops* element, components of a novel system controlling bacterial transcription elongation. *Mol. Microbiol.* 26, 845–851.
- Bailey, M.J., Hughes, C., and Koronakis, V. (2000). *In vitro* recruitment of the RfaH regulatory protein into a specialised transcription complex, directed by the nucleic acid *ops* element. *Mol. Gen. Genet.* 262, 1052–1059.
- Ban, N., Nissen, P., Hansen, J., Moore, P.B., and Steitz, T.A. (2000). The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. *Science* 289, 905–920.
- Belogurov, G.A., Vassilyeva, M.N., Svetlov, V., Klyuyev, S., Grishin, N.V., Vassilyev, D.G., and Artsimovitch, I. (2007). Structural basis for converting a general transcription factor into an operon-specific virulence regulator. *Mol. Cell* 26, 117–129.
- 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.
- 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.
- Ben-Shem, A., Jenner, L., Yusupova, G., and Yusupov, M. (2010). Crystal structure of the eukaryotic ribosome. *Science* 330, 1203–1209.
- Bryan, P.N., and Orban, J. (2010). Proteins that switch folds. *Curr. Opin. Struct. Biol.* 20, 482–488.
- Burmann, B.M., Schweimer, K., Luo, X., Wahl, M.C., Stitt, B.L., Gottesman, M.E., and Rösch, P. (2010). A NusE:NusG complex links transcription and translation. *Science* 328, 501–504.
- Burmann, B.M., Scheckenhofer, U., Schweimer, K., and Rösch, P. (2011). Domain interactions of the transcription-translation coupling factor *Escherichia coli* NusG are intermolecular and transient. *Biochem. J.* 435, 783–789.
- 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.
- 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.
- 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.
- Carter, H.D., Svetlov, V., and Artsimovitch, I. (2004). Highly divergent RfaH orthologs from pathogenic proteobacteria can substitute for *Escherichia coli* RfaH both *in vivo* and *in vitro*. *J. Bacteriol.* 186, 2829–2840.
- Chalissery, J., Muteeb, G., Kalarickal, N.C., Mohan, S., Jisha, V., and Sen, R. (2011). Interaction surface of the transcription terminator Rho required to form a complex with the C-terminal domain of the antiterminator NusG. *J. Mol. Biol.* 405, 49–64.
- 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.
- Fändrich, M., Fletcher, M.A., and Dobson, C.M. (2001). Amyloid fibrils from muscle myoglobin. *Nature* 410, 165–166.
- Greenwald, J., and Riek, R. (2010). Biology of amyloid: structure, function, and regulation. *Structure* 18, 1244–1260.
- Hirtreiter, A., Damsma, G.E., Cheung, A.C.M., Klose, D., Grohmann, D., Vojnic, E., Martin, A.C.R., 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.
- Klein, B.J., Bose, D., Baker, K.J., Yusoff, Z.M., Zhang, X., and Murakami, K.S. (2011). RNA polymerase and transcription elongation factor Spt4/5 complex structure. *Proc. Natl. Acad. Sci. USA* 108, 546–550.
- 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.
- Kyrpides, N.C., Woese, C.R., and Ouzounis, C.A. (1996). KOW: a novel motif linking a bacterial transcription factor with ribosomal proteins. *Trends Biochem. Sci.* 21, 425–426.
- Li, J., Mason, S.W., and Greenblatt, J. (1993). Elongation factor NusG interacts with termination factor ρ to regulate termination and antitermination of transcription. *Genes Dev.* 7, 161–172.
- Littler, D.R., Harrop, S.J., Goodchild, S.C., Phang, J.M., Mynott, A.V., Jiang, L., Valenzuela, S.M., Mazzanti, M., Brown, L.J., Breit, S.N., and Curmi, P.M. (2010). The enigma of the CLIC proteins: Ion channels, redox proteins, enzymes, scaffolding proteins? *FEBS Lett.* 584, 2093–2101.
- Luo, X., Tang, Z., Xia, G., Wassmann, K., Matsumoto, T., Rizo, J., and Yu, H. (2004). The Mad2 spindle checkpoint protein has two distinct natively folded states. *Nat. Struct. Mol. Biol.* 11, 338–345.
- Luo, X., Hsiao, H.H., Bubunenko, M., Weber, G., Court, D.L., Gottesman, M.E., Urlaub, H., and Wahl, M.C. (2008). Structural and functional analysis of the *E. coli* NusB-S10 transcription antitermination complex. *Mol. Cell* 32, 791–802.
- Martinez-Rucobo, F.W., Sainsbury, S., Cheung, A.C., and Cramer, P. (2011). Architecture of the RNA polymerase-Spt4/5 complex and basis of universal transcription processivity. *EMBO J.* 30, 1302–1310.
- Minor, D.L., Jr., and Kim, P.S. (1996). Context-dependent secondary structure formation of a designed protein sequence. *Nature* 380, 730–734.
- Mooney, R.A., Davis, S.E., Peters, J.M., Rowland, J.L., Ansari, A.Z., and Landick, R. (2009a). Regulator trafficking on bacterial transcription units *in vivo*. *Mol. Cell* 33, 97–108.
- Mooney, R.A., Schweimer, K., Rösch, P., Gottesman, M.E., and Landick, R. (2009b). Two structurally independent domains of *E. coli* NusG create regulatory plasticity via distinct interactions with RNA polymerase and regulators. *J. Mol. Biol.* 391, 341–358.
- 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.
- 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.
- Selenko, P., Sprangers, R., Stier, G., Bühler, D., Fischer, U., and Sattler, M. (2001). SMN tudor domain structure and its interaction with the Sm proteins. *Nat. Struct. Biol.* 8, 27–31.
- Sevostyanova, A., Belogurov, G.A., Mooney, R.A., Landick, R., and Artsimovitch, I. (2011). The β subunit gate loop is required for RNA polymerase modification by RfaH and NusG. *Mol. Cell* 43, 253–262.
- Squires, C.L., and Zaporozhets, D. (2000). Proteins shared by the transcription and translation machines. *Annu. Rev. Microbiol.* 54, 775–798.
- 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.
- Sullivan, S.L., Ward, D.F., and Gottesman, M.E. (1992). Effect of *Escherichia coli* nusG function on lambda N-mediated transcription antitermination. *J. Bacteriol.* 174, 1339–1344.

- Surewicz, W., and Apostol, M. (2011). Prion Protein and Its Conformational Conversion: A Structural Perspective. In *Prion Proteins*, J. Tatzelt, ed. (Berlin: Springer), pp. 135–167.
- Tokuriki, N., and Tawfik, D.S. (2009). Protein dynamism and evolvability. *Science* 324, 203–207.
- Tuinstra, R.L., Peterson, F.C., Kutlesa, S., Elgin, E.S., Kron, M.A., and Volkman, B.F. (2008). Interconversion between two unrelated protein folds in the lymphotactin native state. *Proc. Natl. Acad. Sci. USA* 105, 5057–5062.
- Werner, F., and Grohmann, D. (2011). Evolution of multisubunit RNA polymerases in the three domains of life. *Nat. Rev. Microbiol.* 9, 85–98.
- Zhou, H., Liu, Q., Gao, Y., Teng, M., and Niu, L. (2009a). Crystal structure of NusG N-terminal (NGN) domain from *Methanocaldococcus jannaschii* and its interaction with rpoE^{II}. *Proteins* 76, 787–793.
- Zhou, K., Kuo, W.H.W., Fillingham, J., and Greenblatt, J.F. (2009b). Control of transcriptional elongation and cotranscriptional histone modification by the yeast BUR kinase substrate Spt5. *Proc. Natl. Acad. Sci. USA* 106, 6956–6961.