



**HAL**  
open science

## Domain interactions of the transcription:translation coupling factor *E. coli* NusG are intermolecular and transient.

Bjoern Burmann, Ulrich Scheckenhofer, Kristian Schweimer, Paul Rösch

### ► To cite this version:

Bjoern Burmann, Ulrich Scheckenhofer, Kristian Schweimer, Paul Rösch. Domain interactions of the transcription:translation coupling factor *E. coli* NusG are intermolecular and transient.. *Biochemical Journal*, 2011, 435 (3), pp.783-789. 10.1042/BJ20101679 . hal-00586465

**HAL Id: hal-00586465**

**<https://hal.science/hal-00586465>**

Submitted on 16 Apr 2011

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# **Domain interactions of the transcription:translation coupling factor *E. coli* NusG are intermolecular and transient.**

Björn M. Burmann, Ulrich Scheckenhofer, Kristian Schweimer, and Paul Rösch\*

Lehrstuhl für Biopolymere & Research Center for Bio-Macromolecules, Universität Bayreuth,  
Universitätsstraße 30, 95447 Bayreuth, Germany

Running title: NusG domain interactions

Corresponding author; Universität Bayreuth, Lehrstuhl Biopolymere, Universitätsstr. 30,  
95447 Bayreuth, Germany; Phone +49 921 55-3541; Fax: +49 921 553544; E-mail:  
roesch@unibt.de

THIS IS NOT THE VERSION OF RECORD - see doi:10.1042/BJ20101679

Accepted Manuscript

## Synopsis

The bacterial transcription factor NusG is suggested to act as a key coupling factor between transcription and translation (Burmam *et al.* 2010 Science. 328:501-504). and contributes to phage  $\lambda$  mediated antitermination in *E. coli* that enables read-through of early transcription termination sites. *E. coli* NusG consists of two structurally and functionally distinct domains that are connected *via* a flexible linker. The homologous *Aquifex aeolicus* NusG with a secondary structure that is highly similar to *E. coli* NusG shows direct interaction between its N-terminal and its C-terminal domain in a domain-swapped dimer. Here we performed NMR paramagnetic relaxation enhancement measurements and identified interdomain interactions that were concentration dependent and thus likely not only weak and transient, but also predominantly intermolecular. This notion of two virtually independent domains in a monomeric protein was supported by  $^{15}\text{N}$  relaxation measurements. Thus, we suggest a regulatory role of NusG interdomain interactions to be highly unlikely.

## Introduction

*E. coli* NusG (N-utilization substance G) is essential for cell viability [1] and is found in all known bacteria. The C-terminal domain, CTD, contains a KOW motif that is also found in archaeal [2] and eukaryotic [3] proteins. NusG's exact role in transcription regulation remained obscure until very recently when NusG was identified as the long sought-after coupling factor between transcription and translation in *E. coli* [4], in addition to its better known role in the transcription-elongation complex (TEC) where it increases the elongation rate of the RNAP *in vivo* and *in vitro* [5-7] by suppression of transcriptional pausing [7], which could be shown recently to be due to promoting forward translocation of the RNAP [8]. NusG, along with the other Nus factors A, B, E (ribosomal protein S10), RNA, and the RNAP, is also a component of the phage  $\lambda$  protein N mediated antitermination [9-11] and the phage HK022 protein Nun mediated termination complex [12,13]. The regulatory functionality of NusG can be attributed to direct interaction with the RNAP [9,14,15].

Additionally, NusG is involved in transcription antitermination in ribosomal RNA (rRNA)[16,17] and in termination at certain  $\rho$  dependent sites where it recruits  $\rho$  [12,18,19]. NusG also prevents backtracking of the RNAP [7], but detailed information on the final steps of  $\rho$ -dependent termination is still lacking [20]. Direct NusG: $\rho$  interaction was reported [18,19], and the NusG carboxy-terminal domain (CTD) was identified as the interaction domain [4]. In addition to its role in transcription, NusG is also involved in translational regulation [21], and, again, the NusG CTD was identified as the ribosome interaction domain [4].

The structure of *E. coli* NusG [20] is highly similar to the structures of the homologous *Aquifex aeolicus* [22,23] and *Thermus thermophilus* [24] proteins. The *A. aeolicus* NusG is reported to form a domain-swapped dimer under certain crystallization conditions [23]. However, such a dimer is not observed in a different crystallization setup [22], and ultracentrifugation analysis showed no indication of a non-monomeric form of NusG as an isolated protein [19]. Interestingly, however, the NusG paralogue RfaH is reported to exist in a closed conformation with both domains tightly interacting, and this tight interaction is abolished only upon binding of RfaH to the *ops*-site (operon polarity suppressor) of the non-template DNA strand, when the RfaH-CTD leaves the binding pocket on the RfaH-NTD, thus providing an RfaH activation step [25,26]. Although *E. coli* NusG and RfaH exhibit the same folding topology of their NTDs [20,26](NusG-NTD, PDB-ID: 2K06; RfaH, PDB-ID: 2OUG), their CTDs show reversed topologies: Whereas RfaH-CTD is all  $\alpha$ -helical, NusG-CTD exhibits an anti-parallel  $\beta$ -barrel-like structure [20] (NusG-CTD, PDB-ID: 2JVV).

The area masked on the RfaH-NTD by its CTD consists of several surface-exposed hydrophobic residues [15,27], and surface hydrophobic residues are more numerous in RfaH than in NusG [20] where they may be involved in additional protein interactions [22]. These lines of results raises the question which properties the common ancestor of these proteins had, in particular whether it was in closed or open form, and whether domain opening provided a regulatory mechanism. To proceed to an answer to this question, among others, we decided to ask whether *E. coli* NusG domains interact and whether a closed conformation, if only lowly populated, exists.

To address this question and to gain insight into the role of the individual *E. coli* NusG domains within the TEC, their relative motional behavior and their mutual interactions we employed chemical shift and paramagnetic relaxation enhancement (PRE) perturbation titration as well as relaxation experiments.

## Experimental

*Cloning, expression, and purification of full-length NusG and the individual domains:* Cloning, expression and purification was based on published methods [19,20]. NusG was cloned via Bpu1102I and NdeI restriction sites into the *E. coli* expression vector pET11A (Novagen, Madison, WI, USA). *E. coli* strain BL21(DE3) (Novagen, Madison, WI, USA) harboring the recombinant plasmid was grown at 37 °C in LB medium (Luria-Bertani) containing ampicillin (100 µg/ml) until an OD<sub>600</sub> = 0.8 was reached and the cells were induced to a final concentration of 1 mM isopropyl 1-thio-β-D-galactopyranoside (IPTG). Cells were harvested 4 h after induction, resuspended in 4 times the pellet weight of lysis buffer (50 mM TRIS, 150 mM NaCl, 1 mM DTT, pH 7.5, ½ Protease inhibitor tablet [Complete, EDTA-free, Roche Diagnostics GmbH, Mannheim, Germany]), and lysed by using a micro-fluidizer (Microfluidics, Newton, MA, USA). After centrifugation Polyethylenimine (Fluka, Basel, Switzerland) was added drop-wise under continuous stirring to the supernatant to a final concentration of 0.6%. The lysate was incubated for 20 min and centrifuged at 12.000 g. Ammonium sulfate was added drop-wise to a final concentration of 60% under stirring to the supernatant. The lysate was centrifuged at 12.000 g and the pellet was solved in 30 ml of buffer A (50 mM TRIS, pH 7.5). The lysate was afterwards dialyzed against 2 x 4 l buffer A overnight. The lysate was applied to a HeparinFF column (GE Healthcare, Munich, Germany) using a step gradient with increasing NaCl concentrations (0 – 1 M). For further purification the eluted fractions containing NusG were pooled and concentrated with Vivaspin concentrators (Vivascience, MWCO 5 kDa). The concentrated sample was applied to an S75 gel filtration column (50 mM TRIS, 150 mM NaCl; GE Healthcare, Munich, Germany). The fractions containing NusG were pooled and dialyzed against buffer as used for NMR measurements (NMR-buffer: 10 mM potassium phosphate, 50 mM NaCl, pH 6.4). The identity and structural integrity of purified protein was analyzed by SDS-PAGE as well as by NMR spectroscopy. NusG-NTD (1-124) was cloned via BPU1102I and NdeI into the *E. coli* expression vector pET11A (Novagen, Madison, WI, USA). The same procedure as described above for full-length NusG was used for expression and purification of NusG-NTD. NusG-CTD (123-181) was cloned and purified as described before [4].

*Full-length NusG mutation S16C:* For NusG<sup>S16C</sup> the following primers were used: 5'-primer: GTC GTT CAG GCG TTT TGC GGT TTT GAA GGC CGC, 3'-primer: GCG GCC TTC AAA ACC GCA AAA CGC CTG AAC GAC. The mutation was introduced by using the QuikChange protocol (Stratagene, La Jolla, CA, USA). Expression and purification were as for full-length NusG, except that 1 mM DTT was added to all buffers.

*Random spin labeling of the ε-amino groups of lysines:* Spin labeling of the lysines with 1-Oxyl-2,2,5,5-tetramethylpyrrolidine-3-carboxylate N-hydroxysuccinimide Ester (OXYL-1-NHS; Toronto research Chemicals Inc., North York, ON, Canada) was done according to a slightly modified published protocol [28]. Briefly, 500 µM protein solution in NMR buffer (10 mM potassium phosphate, 50 mM NaCl, pH 6.4) were washed with 10 mM sodium carbonate buffer, pH 9.2, for buffer exchange in Vivaspin concentrators (Vivascience, MWCO 5 kDa). A 10 fold excess of OXYL-1-NHS in DMSO was added, followed by incubation for one hour at room temperature and an additional hour at 4 °C in the dark. To remove unreacted OXYL-1-NHS, the sample was washed with 20 ml of NMR-buffer in a Vivaspin concentrator. For reduction of the spin label, ascorbate (500 mM stock) was added directly to the NMR tube to a final concentration of 5 mM.

*Site specific spin labeling of NusG cysteine mutant:* Spin labeling of the cysteines introduced into NusG<sup>S16C</sup> with (1-Oxyl-2,2,5,5-tetramethyl-Δ3-pyrroline-3-methyl) Methanethiosulfonate (MTSL; Toronto research Chemicals Inc., North York, ON, Canada) was done according to a slightly modified published protocol [29]. Briefly, 500 µM protein solution in NMR buffer was washed with 10 mM acetate buffer for buffer exchange in Vivaspin concentrators (Vivascience, MWCO 5 kDa). After addition of DTT to a final concentration of 5 mM, the sample was kept at 4 °C for one hour.

For removal of DTT, the solution was eluted isocratically with 10 mM acetate from a HiTrap Desalting column (GE Healthcare, Munich). A 10 fold excess of MTSL dissolved in acetonitrile was added, followed by incubation overnight at room temperature in the dark. To remove unreacted MTSL, the sample was again washed with 20 ml of NMR-buffer in a Vivaspin concentrator. For reduction of the spin label, ascorbate (500 mM stock) was added directly to the NMR tube to a final concentration of 5 mM.

**NMR:** NMR experiments were recorded at sample temperatures of 298 K on Bruker Avance 600 MHz and 700 MHz (equipped with a cryogenically cooled probe) NMR spectrometers. Assignments for NusG-NTD and full-length NusG were from previous work [20]. The isolated NusG-CTD  $^1\text{H}$ ,  $^{15}\text{N}$  heteronuclear single quantum coherence (HSQC) resonances matched the corresponding signals of full-length NusG perfectly and were trivially assigned and the assignments verified by triple resonance NMR experiments.

Paramagnetic relaxation enhancement ( $R_{2,\text{para}} = R_{2,\text{spinlabel}} - R_{2,\text{no spinlabel}}$ ) was determined by a two point scheme using an HSQC experiment with additional spin echo period during the first INEPT transfer [30]. Spin echo intervals were set to 0.1 ms and 10.2 ms.  $^{15}\text{N}$  longitudinal ( $R_1$ ) and transverse ( $R_2$ ) relaxation rates were determined by standard methods at a  $^1\text{H}$  frequency of 600.2 MHz and 700.2 MHz. Relaxation curves were fitted to monoexponential decays using the program package CURVEFIT (A.G. Palmer, Columbia University, USA). Rotational correlation times were determined from the  $R_2/R_1$  ratios using the program TENSOR2 [31].

**Molecular Dynamics Simulation:** In order to test the possibility of an intramolecular inter-domain interaction of *E. coli* NusG analogous to the swapped dimer interaction in *A. aeolicus* NusG, we calculated the time-dependent structure with molecular dynamics simulations. An initial structural model of *E. coli* NusG with domain interaction was constructed by fitting the solution structures of the *E. coli* domains to *A. aeolicus* NusG. The simulation using XPLOR (XPLOR-NIH 2.1.2 [32]) was based on our standard protocol for protein structure calculation [33]. During the calculation, backbone coordinates were held fixed for all residues except the linker region between NTD and CTD (residues 117-126) and the sequence region 99-104 to reduce steric clashes between NTD in CTD in the initial structure. Additionally, residues in the flexible region E48-R62 were removed *in silicio* to avoid steric conflicts. No further restraints were applied. Existence of an intramolecular inter-domain interaction was accepted based on the absence of violation of restraints (deviation of bond lengths and angles), and 50 structures out of total 80 calculated were accepted based on this criterion.

## Results and Discussion

**NMR titrations suggest independence of the NusG domains:**  $^1\text{H}$ ,  $^{15}\text{N}$ -HSQC chemical shifts of resonances from isolated NusG-CTD and isolated NusG-NTD were virtually identical to the chemical shifts of the respective domains in full-length NusG (Fig. S1), rendering persistent domain interactions within full-length NusG highly unlikely. In order to further clarify whether or not transient domain interactions exist for the more general case of the isolated domains we resorted to observation of HSQC chemical shift perturbations of an isotope labeled domain, NusG-CTD or NusG-NTD, upon addition of the unlabeled potentially interacting domain, a method that is well established as a tool to study the interplay between molecules [34]. NMR spectroscopy, due to its inherent insensitivity and its resulting requirement for sample concentrations in the high micromolar range, is useful for detecting even weak ( $k_d \sim \mu\text{M}$ – $\text{mM}$ ) interactions, and perturbations of amide group resonance shifts as detected by  $^1\text{H}$ ,  $^{15}\text{N}$  HSQC are very sensitive even to subtle structural changes. As NusG-CTD F165 was claimed to strongly interact with a NusG-NTD hydrophobic cavity [23], significant chemical shift changes are expected on domain contact at least for F165 and residues in its vicinity, that is residues in the loop between strands  $\beta_3$  and  $\beta_4$ . Titration of unlabeled NusG-NTD to  $^{15}\text{N}$  labeled NusG-CTD to a two-fold excess and *vice versa*, however, did not result in observable chemical shift changes in the NMR experiments (Fig. S2). Thus, with the level of sensitivity provided by HSQC perturbation experiments, domain interaction could be detected neither within the full-length protein nor for the isolated domains under our experimental conditions. This observation strongly correlates with our earlier conclusion that the area around F165 is not involved

in mutual interactions of NusG domains, but is rather a key residue in NusE:NusG complex formation [4].

*<sup>15</sup>N relaxation reveals decoupled domain reorientation:* In order to study the degree of motional independence of the NusG domains we determined their <sup>15</sup>N relaxation rates in the context of the full-length protein [35]. In a two-domain protein, concerted tumbling of domains can be described by a single rotational diffusion tensor, whereas independent tumbling of domains requires description with different rotational diffusion tensors [36]. In an isotropic rotation model, differences in these tensors are directly reflected in differences in the average relaxation rates.

<sup>15</sup>N transverse and longitudinal relaxation rates were determined at 14.1 T, with a sample concentration of 200 μM to reduce aggregation (Fig. 1A, 1B). In the HSQC spectrum of the full-length protein, virtually all residues of NusG-CTD were observed, the average longitudinal relaxation rate ( $R_1$ ) was  $1.35 \pm 0.09 \text{ sec}^{-1}$ , and the transversal relaxation rate ( $R_2$ ) was  $15.4 \pm 1.9 \text{ sec}^{-1}$ , corresponding to a rotational correlation time of  $\tau_c = 10.1 \text{ ns}$  in an isotropic model. Although the NusG-NTD signals were considerably weaker, 67 non-overlapping signals were observed and the average longitudinal relaxation rate was  $1.1 \pm 0.2 \text{ sec}^{-1}$ , the transversal rate  $19.0 \pm 4.6 \text{ sec}^{-1}$ , corresponding to  $\tau_c = 13.1 \text{ ns}$  in the isotropic model. Analysis of  $R_1/R_2$  distributions offers an elegant method to detect interdomain motion on a time scale faster than the overall tumbling [37]. The  $R_1/R_2$  ratios form the basis for determination of the rotational diffusion tensor by NMR relaxation, and for compact globular proteins a uniform distribution of  $R_1/R_2$  ratios is characteristic [35]. The bimodal distribution of the  $R_1/R_2$  ratios (Fig. 1C) that reflect the two domains demonstrates their different rotational reorientation behavior, arguing against a stable domain interaction on the nanosecond timescale. Although the difference in relaxation rates demonstrates independent movement of the two domains to a certain degree,  $\tau_c$  of the NTD and CTD in full-length NusG are higher than  $\tau_c$  of the isolated domains (13.1 ns vs. 8.9 ns and 10.1 ns vs. 4 ns; see Fig. S3 for relaxation rates of the individual domains). This increase of rotational correlation times of individual isolated domains up to values found in multidomain proteins without any domain interaction has been observed for several proteins [38-41]. The spatial exclusion due to the presence of a second domain connected by a linker is likely hampering an unrestricted rotation of a domain resulting in a longer correlation time [39,41]. The large increase of the rotational correlation times of both NusG domains indicates that motional decoupling of NusG-CTD and NusG-NTD via the 5 residue linker is imperfect, and a significant contribution of this effect in full length NusG to the relaxation mechanisms is expected. The interdomain motion alters the overall shape of full length NusG and therefore affects the overall tumbling. As consequence, complete uncoupling of domain dynamics and overall rotation is unlikely, and this coupled motions render a more detailed analysis at least very difficult [39]. In intact NusG, V162 and I164 of NusG-CTD exhibit significantly enhanced  $R_2$  rates of  $21.5 \text{ sec}^{-1}$  as compared to the domain average of  $15.4 \text{ sec}^{-1}$ , a difference not found in isolated CTD. Most likely the enhanced transverse relaxation can be attributed to a chemical exchange contribution, and as these residues are located in the domain interface of the swapped-dimer crystal structure, this might be regarded as initial evidence of transient domain interaction in solution.

*Only highly sensitive PRE experiments show signs of interaction between the isolated domains:* Fast exchange on the NMR timescale leads to observation of population averaged parameters, and states that are extremely weakly populated on time average often cannot be detected even by HSQC chemical shift perturbations or changes in relaxation rates. Paramagnetic interactions, however, provide a means to detect even weak and transient interactions between molecules. Paramagnetic centers dramatically increase relaxation rates of nearby nuclei and, as a consequence, presence of even very minor concentrations of paramagnetic labels in proximity to observed nuclei can enhance the relaxation rate of the latter to an observable degree [42]. Thus, PRE observed in a non spin-labeled protein in the presence of a spin-labeled protein points to an at least transient proximity of both molecules. The experiment is straightforward as a non-spin-labeled protein, for example NusG-CTD, can be made easily detectable in HSQC spectra by <sup>15</sup>N enrichment, and introduction of paramagnetic centers into proteins, for example NusG-NTD, is possible by random labeling of mostly surface exposed lysine residues with OXYL-1-NHS [43]. In such an experiment, addition of OXYL-1-NHS-NusG-NTD to <sup>15</sup>N-NusG-CTD (Fig. S4) causes observable and specific PRE in the latter, and *vice versa* (Fig. 2). Mapping of the resonances with increased relaxation rates onto the three-

dimensional structure of NusG-CTD shows very clearly the region around F165 and the region P140 to N145, close to F165 in the turn between  $\beta^1$  and  $\beta^2$ , to be the ones most seriously affected (Fig. 2).

The outcome of the inverse experiment, that is addition of OXYL-1-NHS-NusG-CTD to  $^{15}\text{N}$ -NusG-NTD (Fig. S4), was less clear-cut. The hydrophobic patch around F65 was clearly affected, together with several residues in helix  $\alpha^3$  and residues close to F65 that form an apolar surface surrounded by polar residues. However, several isolated residues in helices  $\alpha^1$  and  $\alpha^2$  at the opposite side of the molecule as well as several residues of the »upper« loop regions facing away from the putative interaction site were affected (Fig. 2). Combined, the PRE results yield a picture in which a plug around F165 fits nicely into a socket around F65. Although this is a reasonable contact surface, and although these sites correspond well to those found in the swapped dimer crystal structure [23], our negative results with HSQC chemical shift mapping indicate the interaction to be rather weak and transient. Under the assumption of chemical shift changes in the range typically observed in this experimental setup (e.g. NusG-CTD:NusE [4],  $\lambda\text{N:AR1}$  [44], LckSH3:Tip [33]) no significant chemical shift changes point to less than 10 % bound species in the NMR samples, which corresponds to a  $K_d$  higher than 5 mM.

*PREs of site-specific full-length NusG are concentration dependent:* Chemically modifying cysteines with MTSL provides a means to observe the PRE induced by site-specific labels. Thus, in order to clarify whether or not the weak domain interaction observed by random labeling in solution is intra- or intermolecular, we performed additional PRE experiments on a full-length NusG construct, containing MTSL-cysteine S16C. PRE originating from MTSL-Cys was detected in the unchanged CTD, clearly indicating an interaction between NusG-NTD and NusG-CTD, thus neither supporting nor ruling out the formation of a domain-swapped dimer or an intramolecular inter-domain interaction within one molecule. To further probe these domain interactions, we studied the concentration dependence of the PRE induced by MTSL-Cys16 on the CTD as intramolecular domain interactions would clearly be expected to result in a concentration-independent PRE due to unchanged internuclear distance distributions in a stable structure with interacting domains, whereas a concentration dependent PRE would demonstrate intermolecular domain interactions by reduction of the fraction of interacting molecules upon dilution and decrease of the observed ensemble averaged PRE. The PRE thus determined at different protein concentrations (Fig. 3) shows a direct dependence of the relaxation enhancement on sample concentration, thus clearly out-ruling an intramolecular interaction and deciding for the model of a transient domain-swapped dimer formation, which is extremely scarcely populated.

*E. coli* NusG NTD:CTD interaction – analogous to that observed within the swapped dimer in *A. aeolicus* NusG, but within a single molecule, would only be possible if no sterical restraints would exist in the linker sequence. To test whether or not such steric and dynamic restraints exist we performed molecular dynamics simulations using a structural model with interacting domains to explore the conformational space of the linker region. Only a slight rearrangement of the loop prior to helix  $\alpha^3$  of NusG-NTD would be necessary to remove all sterical clashes in the initial model based on the solution structure of *E. coli* NusG. In addition, all 50 analyzed lowest energy structures are without violation of bonds or angles and provide possible linker conformations that could allow domain interactions within one NusG molecule (Fig. 4). Therefore, an intramolecular domain interaction cannot be ruled out solely on the argument of the linker being too short. Nevertheless, the linker region must adopt a very extended conformation for the domains to be able to directly interact. This is expected to lead to a high degree of rigidity in the linker on the timescale of molecular reorientation due to the low number of possible conformational states that the linker can adopt in this extended state. On the other hand, the  $^{15}\text{N}$  transversal relaxation rates of G119 (7.6 Hz) and D120 (6.4 Hz) are significantly lower than those for the other residues in the respective domains (Fig. 1B). The reduction of  $R_2$  is characteristic for enhancement of mobility on the sub- $\tau_c$  (sub nanosecond) timescale, demonstrating experimentally the flexibility of the linker region in contrast to the rigidity of the linker needed for direct interaction as shown by the MD simulation. This renders a domain interaction within one molecule highly unlikely.

**Conclusions:** Our NMR studies on domain interactions of *E. coli* NusG show these interactions to be not detectable *via* standard chemical shift mapping, and  $^{15}\text{N}$ -relaxation measurements reveal

movements of the two domains to be independent to a very high degree. Domain interactions could only be observed with the extremely sensitive technique of PRE measurements, and it needed a combination of this technique and site-directed paramagnetic labeling to show that these weak interactions were inter- rather than intramolecular. Thus, a low population of molecules seems to adopt a bimolecular conformation compatible with the proposed swapped dimer observed in X-ray crystallography. Intramolecular domain interaction, however, would be required for an autoinhibitory function and therefore biologically relevant form to reduce the population of the active open state. The mode of regulation found for the NusG paralogue RfaH [26] can thus be ruled out to work for NusG, which may explain why NusG maintains its overall function without additional activation signals, while RfaH needs the *ops*-site to render the protein functional. Current opinion is that NusG and RfaH derived from the same ancestor [15], possibly a NusG protein that exists in closed conformation and that may be regulated by domain interactions. It may further be speculated that regulation of NusG *via* domain interactions is no longer necessary when RfaH is present. This hypothesis awaits further structural investigation, because in addition to the proposed swapped-dimer for *A. aeolicus* [23] and the non-interacting *T. thermophilus* [24] NusGs, detailed structural information on other thermophilic NusGs is not available yet.

### Acknowledgement

We thank Ramona Heissmann for excellent technical assistance.

### Funding

This project was supported by a grant to PR from the Deutsche Forschungsgemeinschaft DFG (Ro617/16-1)



## REFERENCES

- 1 Downing, W. L., Sullivan, S. L., Gottesman, M. E. and Dennis, P. P. (1990) Sequence and transcriptional pattern of the essential *escherichia coli* *secE-nusG* operon. *J. Bacteriol.* **172**, 1621-1627.
- 2 Kyrpides, N. C. and Ouzounis, C. A. (1999) Transcription in archaea. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 8545-8550.
- 3 Hartzog, G. A., Wada, T., Handa, H. and Winston, F. (1998) Evidence that Spt4, Spt5, and Spt6 control transcription elongation by RNA polymerase II in *saccharomyces cerevisiae*. *Genes Dev.* **12**, 357-369.
- 4 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.
- 5 Burova, E., Hung, S., Sagitov, V., Stitt, B. and Gottesman, M. (1995) *Escherichia coli* NusG protein stimulates transcription elongation rates in vivo and in vitro. *J. Bacteriol.* **177**, 1388-1392.
- 6 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.
- 7 Artsimovitch, I. and Landick, R. (2000) Pausing by bacterial RNA polymerase is mediated by mechanistically distinct classes of signals. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 7090-7095
- 8 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.
- 9 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.
- 10 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.
- 11 Mason, S. W., Li, J. and Greenblatt, J. (1992) Host factor requirements for processive antitermination of transcription and suppression of pausing by the N protein of bacteriophage lambda. *J. Biol. Chem.* **267**, 19418-19426.
- 12 Sullivan, S. L. and Gottesman, M. E. (1992) Requirement for *E. coli* NusG protein in factor-dependent transcription termination. *Cell.* **68**, 989-994.
- 13 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.
- 14 Mason, S. W. and Greenblatt, J. (1991) Assembly of transcription elongation complexes containing the N protein of phage lambda and the *escherichia coli* elongation factors NusA, NusB, NusG, and S10. *Genes Dev.* **5**, 1504-1512.

- 15 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.
- 16 Torres, M., Balada, J. M., Zellars, M., Squires, C. and Squires, C. L. (2004) In vivo effect of NusB and NusG on rRNA transcription antitermination. *J Bacteriol.* **186**, 1304-1310.
- 17 Zhou, Y., Filter, J. J., Court, D. L., Gottesman, M. E. and Friedman, D. I. (2002) Requirement for NusG for transcription antitermination in vivo by the lambda N protein. *J Bacteriol.* **184**, 3416-3418.
- 18 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.
- 19 Paskan, 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.
- 20 Mooney, R. A., Schweimer, K., Rösch, P., Gottesman, M. E. 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.
- 21 Zellars, M. and Squires, C. L. (1999) Antiterminator-dependent modulation of transcription elongation rates by NusB and NusG. *Mol Microbiol.* **32**, 1296-1304.
- 22 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.
- 23 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.
- 24 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.
- 25 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.
- 26 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.
- 27 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. Nat. Acad. Sci. USA.* **105**, 865-870.
- 28 Lawrence, J. J., Berne, L., Ouvrier-Buffet, J. L. and Piette, L. H. (1980) Spin-label study of histone H1-DNA interaction. comparative properties of the central part of the molecule and the N and C-amino tails. *Eur. J. Biochem.* **107**, 263-269.
- 29 Berliner, L. J., Grunwald, J., Hankovszky, H. O. and Hideg, K. (1982) A novel reversible thiol-specific spin label: Papain active site labeling and inhibition. *Anal Biochem.* **119**, 450-455.

- 30 Iwahara, J., Tang, C. and Marius Clore, G. (2007) Practical aspects of  $(1)H$  transverse paramagnetic relaxation enhancement measurements on macromolecules. *J. Magn. Reson.* **184**, 185-195.
- 31 Dosset, P., Hus, J. C., Blackledge, M. and Marion, D. (2000) Efficient analysis of macromolecular rotational diffusion from heteronuclear relaxation data. *J. Biomol. NMR.* **16**, 23-28.
- 32 Schwieters, C. D., Kuszewski, J. J., Tjandra, N. and Clore, G. M. (2003) The xplor-NIH NMR molecular structure determination package. *J. Magn. Reson.* **160**, 66-74.
- 33 Schweimer, K., Hoffmann, S., Bauer, F., Friedrich, U., Kardinal, C., Feller, S. M., Biesinger, B. and Sticht, H. (2002) Structural investigation of the binding of a herpesviral protein to the SH3 domain of tyrosine kinase Ick. *Biochemistry.* **41**, 5120-5130.
- 34 Zuiderweg, E. R. (2002) Mapping protein-protein interactions in solution by NMR spectroscopy. *Biochemistry.* **41**, 1-7.
- 35 Kay, L. E., Torchia, D. A. and Bax, A. (1989) Backbone dynamics of proteins as studied by  $15N$  inverse detected heteronuclear NMR spectroscopy: Application to staphylococcal nuclease. *Biochemistry.* **28**, 8972-8979.
- 36 Fushman, D., Varadan, R., Assfalg, M. and Walker, O. (2004) Determining domain orientation in macromolecules by using spin-relaxation and residual dipolar coupling measurements. *Prog. NMR Spect.* **44**, 189-214.
- 37 Horstmann, M., Ehses, P., Schweimer, K., Steinert, M., Kamphausen, T., Fischer, G., Hacker, J., Rösch, P. and Faber, C. (2006) Domain motions of the mip protein from legionella pneumophila. *Biochemistry.* **45**, 12303-12311.
- 38 Zhou, H., McEvoy, M. M., Lowry, D. F., Swanson, R. V., Simon, M. I. and Dahlquist, F. W. (1996) Phosphotransfer and CheY-binding domains of the histidine autokinase CheA are joined by a flexible linker. *Biochemistry.* **35**, 433-443.
- 39 Nakaseko, Y., Neuhaus, D., Klug, A. and Rhodes, D. (1992) Adjacent zinc-finger motifs in multiple zinc-finger peptides from SWI5 form structurally independent, flexibly linked domains. *J. Mol. Biol.* **228**, 619-636.
- 40 Konrat, R., Krautler, B., Weiskirchen, R. and Bister, K. (1998) Structure of cysteine- and glycine-rich protein CRP2. backbone dynamics reveal motional freedom and independent spatial orientation of the lim domains. *J. Biol. Chem.* **273**, 23233-23240.
- 41 Beglova, N., North, C. L. and Blacklow, S. C. (2001) Backbone dynamics of a module pair from the ligand-binding domain of the LDL receptor. *Biochemistry.* **40**, 2808-2815.
- 42 Tang, C., Iwahara, J. and Clore, G. M. (2006) Visualization of transient encounter complexes in protein-protein association. *Nature.* **444**, 383-386.
- 43 Jahnke, W., Rudisser, S. and Zurini, M. (2001) Spin label enhanced NMR screening. *J. Am. Chem. Soc.* **123**, 3149-3150.
- 44 Prasch, S., Schwarz, S., Eisenmann, A., Wöhr, B. M., Schweimer, K. and Rösch, P. (2006) Interaction of the intrinsically unstructured phage lambda N protein with *E. coli* NusA. *Biochemistry.* **45**, 4542-4549.

**Figure Legends:**

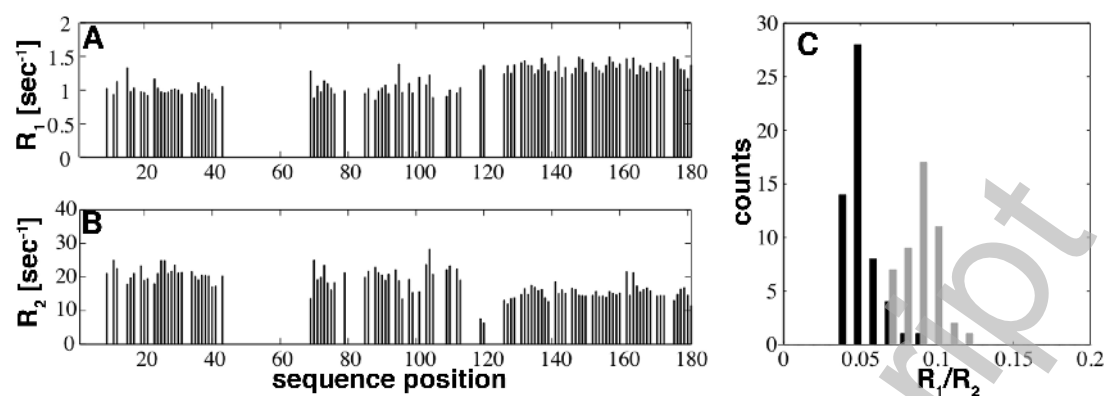
**Fig. 1** A) Longitudinal relaxation rates ( $R_1$ ) for full-length NusG. B) Transversal relaxation rates ( $R_2$ ) for full-length NusG. Residues 45 to 65 and 115 to 125 are highly flexible and located in strongly overlapping regions of the protein so that their relaxation rates could not be determined. C) The distribution of  $R_1/R_2$  is bimodal for full-length NusG. NusG-CTD, black bars; NusG-NTD, gray bars.

**Fig. 2** A) Surface representation of the individual *E. coli* NusG domains (NusG-NTD PDB-ID: 2K06; NusG-CTD PDB-ID: 2JVV; [20]) with secondary structure elements. F65 and F165, blue sticks; amino acids with an  $R_{2,para}(H^N)$ -effect  $>20$  Hz in the titrations of the individual domains, red. B)  $R_{2,para}(H^N)$ -rates for each amino acid upon titration with the individual spin-labeled domains, gray; signals disappearing due to extensive line broadening, red; 20 Hz significance level, dotted line.

**Fig. 3** Effects of spin-labeled NusG<sup>S16C</sup> on the  $R_{2,para}(H^N)$ -rates of NusG-CTD in full-length NusG. Effects at a protein concentration of 150  $\mu$ M on selected amino acids, dark gray; effects at a protein concentration of 50  $\mu$ M, light gray.

**Fig. 4** Superposition of the 20 lowest energy structures of molecular dynamics simulation with intramolecular inter-domain interactions (NTD, green; CTD, red; linker region (residues Q117-T126), blue). The flexible region E48-R62 (indicated by the dashed line in the left picture) was removed *in silicio* to avoid steric conflicts.

Figure 1:



Accepted Manuscript

Figure 2:

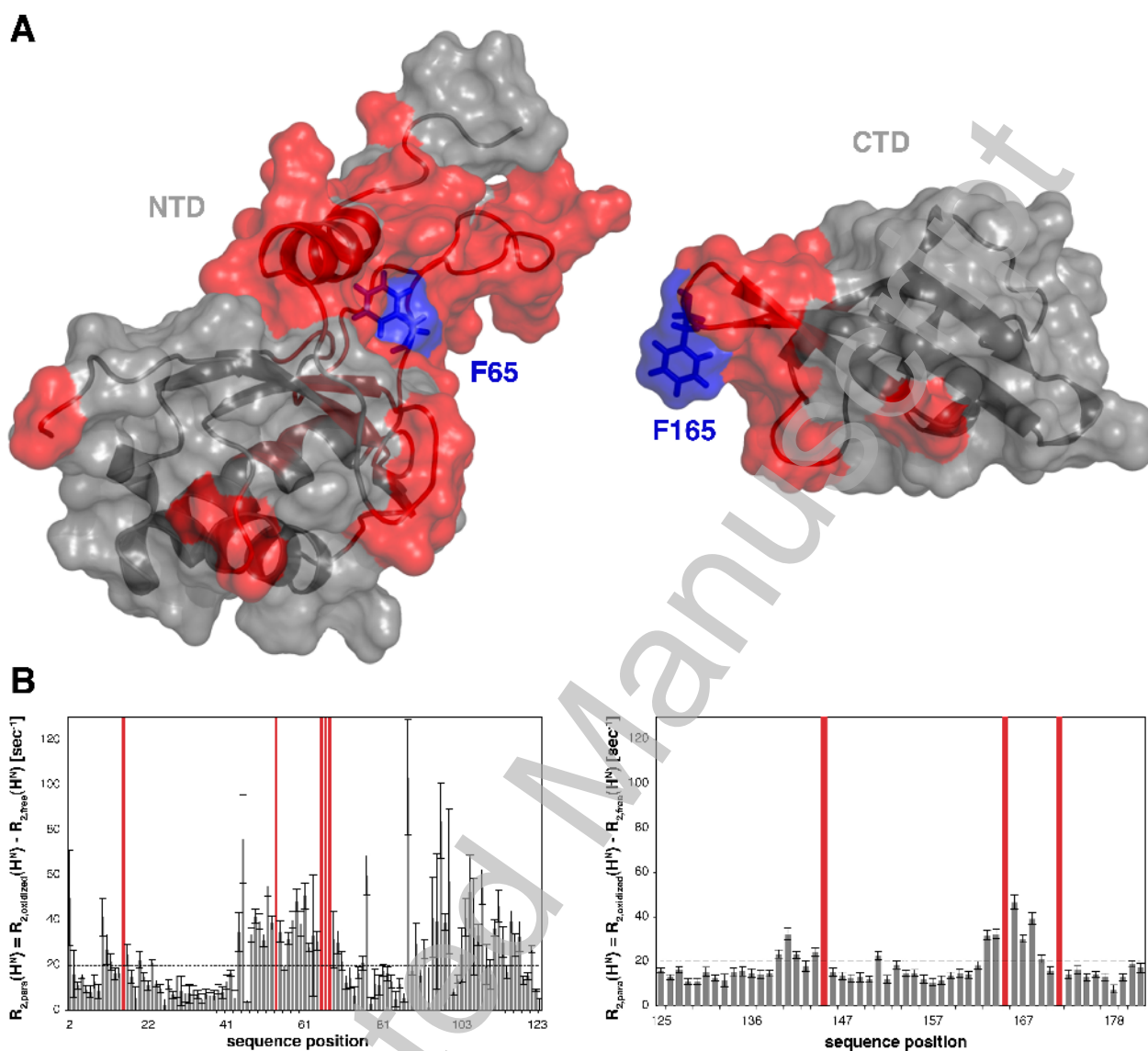
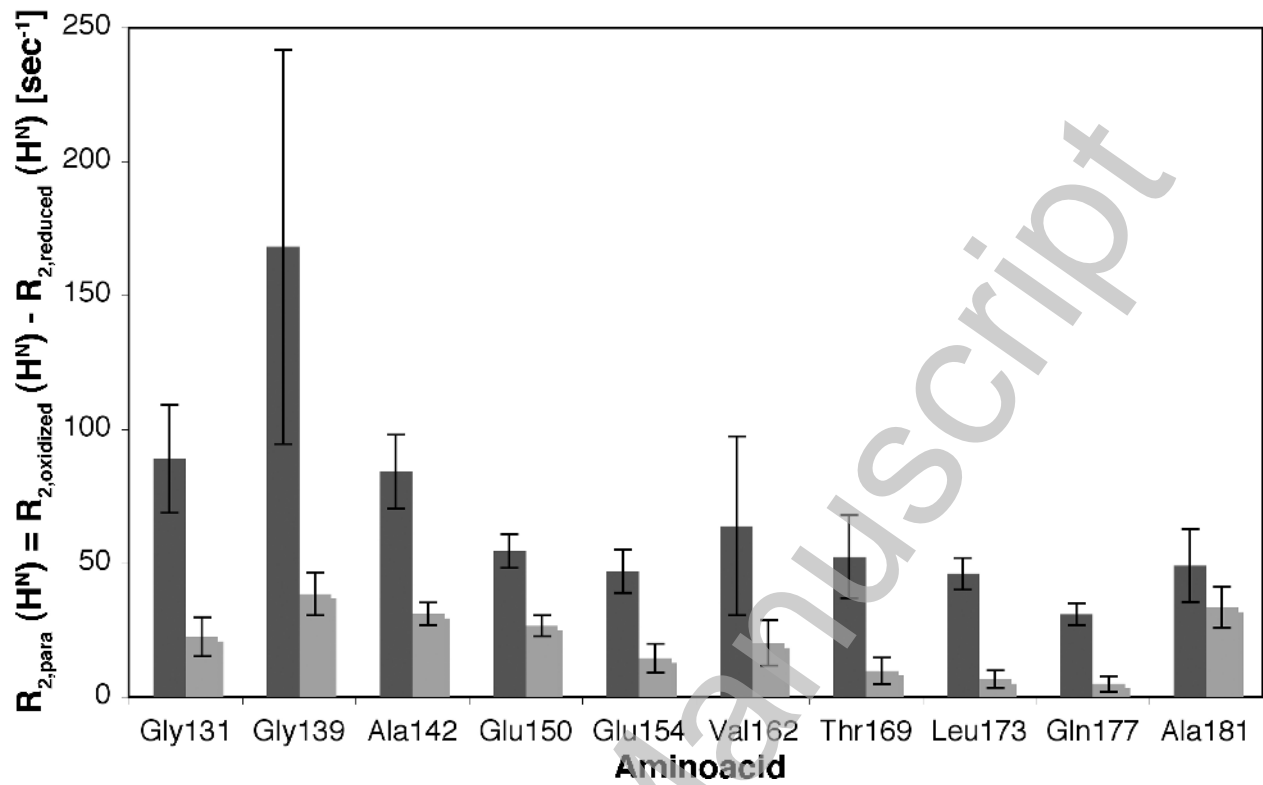


Figure 3:



Accepted Manuscript

Fig. 4:

