Spt4/5 stimulates transcription elongation through the RNA polymerase clamp coiled-coil motif

Angela Hirtreiter1, Gerke E. Damsma2, Alan C. M. Cheung2, Daniel Klose1, Dina Grohmann1, Erika Vojnic2, Andrew C. R. Martin1, Patrick Cramer2,* and Finn Werner1,*

1Division of Biosciences, University College London, Institute for Structural and Molecular Biology, Darwin Building, Gower Street, London WC1E 6BT, UK and 2Department of Biochemistry, Gene Center and Center for Integrated Protein Science Munich, Ludwig-Maximilians-Universität München, Feodor-Lynen-Strasse 25, 81377 Munich, Germany

Received January 15, 2010; Revised February 12, 2010; Accepted February 16, 2010

Abstract

Spt5 is the only known RNA polymerase-associated factor that is conserved in all three domains of life. We have solved the structure of the Methanococcus jannaschii Spt4/5 complex by X-ray crystallography, and characterized its function and interaction with the archaeal RNAP in a wholly recombinant in vitro transcription system. Archaeal Spt4 and Spt5 form a stable complex that associates with RNAP independently of the DNA–RNA scaffold of the elongation complex. The association of Spt4/5 with RNAP results in a stimulation of transcription processivity, both in the absence and the presence of the non-template strand. A domain deletion analysis reveals the molecular anatomy of Spt4/5—the Spt5 Nus-G N-terminal (NGN) domain is the effector domain of the complex that both mediates the interaction with RNAP and is essential for its elongation activity. Using a mutagenesis approach, we have identified a hydrophobic pocket on the Spt5 NGN domain as binding site for RNAP, and reciprocally the RNAP clamp coiled-coil motif as binding site for Spt4/5.

Introduction

All cellular RNA polymerases (RNAPs) from the three domains of life have common structural and functional features, and contain genuine homologues of all bacterial RNAP subunits (1). The archaeal and eukaryotic RNAPs furthermore share a subset of subunits that are not conserved in the bacterial enzyme and two of them, F/E, play a role for the processivity of transcription (1,2). Transcription elongation by all multisubunit RNAP is a discontinuous process that is frequently interrupted by pausing, stalling or arrest (3). This can be detrimental for successful RNA synthesis but also productive in the context of the regulation of gene expression. In eukaryotes, nucleosomes represent an impediment to efficient transcription elongation and a large array of factors influence the progression of RNAP through chromatin by remodelling its structure, but without necessarily interacting with the RNAP itself (4,5). A plethora of elongation factors regulate transcription elongation by interacting directly with RNAP; these include transcript cleavage factors such as eukaryotic TFIIS that increases processivity by releasing arrested elongation complexes, and Spt4/5, which mediates both positive and negative regulation of transcription elongation (5,6).

Spt5 is the only known RNAP-associated transcription factor that is universally conserved in evolution (7); two domains of eukaryotic Spt5 are homologous to the bacterial N-utilization substance G (NusG) factor: the NusG N-terminal domain (NGN) and C-terminal Kyprides–Ouzounis–Woese (KOW) domain (Figure 1E) (8,9). Eukaryotic Spt5 variants include four to six copies of the KOW domain and two C-terminal hepta- and octapeptide repeats (ctr-1 and -2) that are subject to posttranslational phosphorylation (Figure 1E) (10,11). Eukaryotic Spt4/5 variants include four to six copies of the KOW domain and two C-terminal hepta- and octapeptide repeats (ctr-1 and -2) that are subject to posttranslational phosphorylation (Figure 1E) (10,11). Eukaryotic Spt4/5 has both positive and negative effects on transcription. Human Spt4/5 (also called DSIF for DRB-sensitivity inducing sensitivity factor) stimulates the processivity of RNAPII in a purified system (12); it is also required for promoter proximal stalling of RNAPII site (13,14). In Bacteria, NusG has pleiotropic effects on transcription elongation, and termination. NusG increases

*To whom correspondence should be addressed. Tel: +44 20 7679 0147; Fax: +44 (0)20 7679 7193; Email: werner@biochem.ucl.ac.uk

Correspondence may also be addressed to Patrick Cramer. Tel: +49 89 2180 76951; Fax: (+49) 89 2180 76999; Email: cramer@LMB.uni-muenchen.de

© The Author(s) 2010. Published by Oxford University Press.
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/2.5), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.
Figure 1. Structure of the archaeal transcription elongation factor Spt4/5. X-ray structure of *M. jannaschii* Spt4/5NGN complex (A). The interface between Spt4 and Spt5 (B). Alignment of Spt4 and Spt5 NGN homologs. Highly conserved residues are depicted in bold. The Spt5 residues that are mutated are indicated by asterisk. Mj, *Methanocaldococcus jannaschii*; Sc, *Saccharomyces cerevisiae*; Hs, *Homo sapiens*; Ec, *Escherichia coli* (C). Structural comparison of bacterial NusG NGN from *E. coli* (left, pdb 2K06), the archaeal structure of *M. jannaschii* Spt4/5 NGN complex (middle) and eukaryotic Spt4/5 NGN from *S. cerevisiae* (right, pdb 2EXU) (D). Domain architecture of Spt4/5 complexes in the three domains of life (E). A model of the complete *M. jannaschii* Spt4/5 complex (F).
the transcription elongation rate by suppressing pausing and backtracking of RNAP (15–17). NusG enhances rho-dependent termination but in combination with other factors (including NusA, B and E) it binds to RNAP and forms the anti-termination complex that resists pausing and termination (18,19). Despite considerable efforts, the molecular mechanisms by which NusG modulates RNAP are still not completely understood. The NGN domains of NusG and its parologue RfαH have been shown to be necessary and sufficient for the stimulation of transcription elongation (20,21), whereas the KOW domain has been speculated to facilitate interactions with additional factors (rho, NusA, B and E) and nucleic acids (22).

Like bacterial NusG, archaeal Spt5 only contains one KOW domain and lacks the C-terminal repeat (ctr) motifs (7). Like eukaryotic Spt5, archaeal Spt5 forms a complex with Spt4, which has no homologues in bacteria (8). Therefore, archaeal Spt4/5 (a small heterodimeric factor) is a chimera of bacterial and eukaryotic features that is likely to reflect the evolutionary transition state between a relatively simple, small single subunit factor (bacterial NusG) and a large highly regulated heterodimeric complex (eukaryotic Spt4/5, Figure 1E). The evolutionary conservation of Spt5/NusG leads to the intriguing speculation that the molecular mechanisms by which it regulates RNAP are evolutionary ancient, and that a NusG-like precursor already was present in the Last Universal Common Ancestor (LUCA).

Here, we report a structure–function analysis of archaeal Spt4/5. We have solved the crystal structure of the Spt4/5 core from the hyperthermophilic archaeon *Methanocaldococcus jannaschii*. To investigate Spt4/5 function, we used our recombinant archaeal in vitro transcription system (2,23–29). We have identified the minimal domain configuration of archaeal Spt4/5 that is required for the stimulation of elongation and the binding to RNAP, and have identified the regions on Spt5 and RNAP that mediate the interaction. Our results suggest that Spt4/5 structure and function are universally conserved in all three domains of life.

**MATERIALS AND METHODS**

**Crystallization and structure determination of *M. jannaschii* Spt4/5 NGN complex**

The purified core Spt4/5 from *M. jannaschii* was concentrated to 30 mg/ml and crystallization screens were performed at the crystallization facility of the Max Planck Institute of Biochemistry (Martinsried, Germany). Crystals hits were optimized by fine grid screening at 22°C with the hanging drop vapour diffusion method by mixing 1 μl of sample solution with 1 μl of reservoir solution, with the largest crystals appearing in 2–3 days in 50 mM MES pH 5.5–7.0, 20–30% PEG400. Crystals were flash-cooled by plunging into liquid nitrogen and diffraction data collected at beamline X06SA of the Swiss Light Source using a Pilatus 6M pixel detector (30). Raw data were processed with XDS. The structure was solved by molecular replacement using PHASER (31) with the structure of the core Spt5 from *M. jannaschii* (32) together with Spt4 homolog 1RYQ (46% sequence identity). The structure was built and refined using the programs COOT (33), REFMAC (34) and phenix.refine (35).

**Comparative modelling**

The Spt 5 C-terminal domain was modelled based on the crystal structure of *Aquifex aeolicus* NusG (pdb 1M1G) (22) using Modeller (36). Structural alignment in VMD (37) of the Spt 5 N-terminal domain with the respective N-terminus of *A. aeolicus* provided an initial orientation for the Spt 5 C-terminus, which then served as the template. Stereochemistry of the resulting model was assessed using Procheck v. 3.4. All protein structure images were generated using MacPyMol.

**Transcription elongation assay**

Transcription assays were carried out as described in (2) using the parental DNA TS and NTS and RNA primer. The template DNA strand (TS) and the 14 nt RNA primer were annealed in a 1:5 molar ratio for 1 min at 96°C and slowly cooled to ambient temperature. A RNAP were pre-incubated with RNA/DNA scaffold at 65°C in 0.5 x HNME buffer (20 mM HEPES, pH 7.3, 1.25 mM sodium chloride, 1.75 mM magnesium chloride, 0.05 mM EDTA, 2.5% glycerol and 10 mM DTT). To prevent non-specific binding of RNAP to nucleic acid scaffolds and multiround transcription by reinitiation the reaction mix was incubated for further 10 min in presence of 0.05 mg/ml Heparin at 65°C. Transcription reaction was started by addition of nucleotides (750 μM ATP/UTP/GTP, 2 μM CTP, [α-32P] CTP (5000 Ci/mmol), in presence or absence of 2.5 μM NTS. The radiolabelled transcripts were separated on 16% denaturing PAGE (8M Urea, 0.5 x TBE), dried for 1 h at 80°C under vacuum and visualized using an Fuji FLA 2000 reader and Image Analyser software (Fuji, Japan).

**Recombinant proteins**

RNAP were assembled as described previously (28). Expression vectors for GST-fusion proteins of *M. jannaschii* Spt4/5, Spt5 NGN (1-82), Spt5 KOW (83–147) and Spt4 were generated by PCR amplification from genomic DNA and subcloning into BamHI (Spt5) and XhoI (Spt4) sites of pGEX-2TK (GE Healthcare). Spt4 was subcloned into pET151/D-TOPO (Invitrogen) and HisTRAP, Invitrogen). Following cleavage of the GST-tag overnight with thrombin at 4°C, GST was removed by heat treatment (30 min at 65°C) and proteins were purified by size exclusion chromatography (HiPrep-Sepharyl, S100HR 16/60 and S200 HR 26/60, GE Healthcare).
Isotopic labelling of recombinant proteins

Spt4/5, Spt4/5 NGN, Spt5 and Spt5 NGN (100 μg) were radio labelled using gamma 32P-ATP (Perkin Elmer) and cAMP-dependent catalytic subunit PKA (New England Biolabs) for 2 h at 37°C. This protocol generated N-terminally 32P-labelled recombinant proteins via the RRASV kinase recognition site in the vector linker of pGEX-2TK vector. The His-Spt4 protein expressed from pET151D was not suitable for labelling using this approach.

Heat stability assay

Spt4/5 variants were incubated in P300 buffer for 20 min at 65°C or 75°C. Precipitated proteins were removed by centrifugation at 14 000g for 10 minutes and supernatants were subjected to SDS–PAGE.

Transcription elongation assay

Transcription elongation assays were carried out as described previously (2).

RNAP binding assay

Radiolabelled Spt4/5, Spt4/5 NGN, Spt5 and Spt5 NGN (50 nM) were incubated with RNAP (250 and 500 nM) in 0.5x HNME buffer (29) in a total volume of 15 μl at 65°C for 20 min. Following the addition of 5 μl native PAGE loading buffer the reaction was separated on 10% native Tris-Glycine PAGE and the 32P signal was visualized using a phosphor imaging plate and Image Analyser (Fuji FLA 2000, Japan). The mobility of the labelled Spt4/5 variants varies due to their different size and charge. Spt5 forms a doublet that is likely to be due to the dimerization of Spt5 in the absence of Spt4 (32).

RESULTS

Structure of M. jannaschii Spt4/5

An apparently stoichiometric heterodimer of the Spt4 and the Spt5 NGN domain from M. jannaschii was purified after co-expression of the two proteins in bacteria (‘Materials and Methods’ section). The heterodimer was crystallized by vapour diffusion and the structure was solved by molecular replacement using the structure of core Spt5 from M. jannaschii (32) and a Spt4 homologue from Pyrococcus furiosis (1RYQ, ref http://www.ncbi.nlm.nih.gov/pubmed/15858261) as search models (Table 1 and ‘Materials and Methods’ section). Four copies of each search model were located in the asymmetric unit. The structure was built and refined to a resolution of 1.9 Å (Table 1). The presence of Zinc ions in the crystals was confirmed by an X-ray absorption edge scan (Supplementary Figure S6).

The Spt5 NGN domain consists of a four-stranded anti-parallel β-sheet as part of an alpha–beta sandwich (Figure 1A). Spt4 also contains a four-stranded β-sheet stabilized by a zinc ion that is coordinated by the side chains of four conserved cysteine residues (Figure 1A). Mutations altering any of the four zinc-binding cysteines in Saccharomyces cerevisiae Spt4 cause loss-of-function phenotypes and in several instances greatly decrease cellular levels of Spt4 (38,39), suggesting that the zinc-binding site is an essential structural element of Spt4 (8). In the Spt4/5 NGN heterodimer, the β-sheets of Spt4 and Spt5 NGN are aligned edge to edge, forming a continuous eight-stranded anti-parallel β-sheet (Figure 1A). In addition to this β sheet interface, several hydrogen bonds contribute to the specificity of Spt4–Spt5NGN interactions. In particular, Ser39 from Spt4 forms a conserved hydrogen-bonding interaction with Glu46 in Spt5NGN (Figure 1B). This hydrogen bond fixes Glu46, which results in a strong acid–base interaction (Guo et al., 2008; Nicholson et al., 1988).

Additional Spt4-Spt5 interactions are formed by helix a1 and ‘Materials and Methods’ section) as search models (Table 1 and ‘Materials and Methods’ section). Four copies of each search model were located in the asymmetric unit. The structure was built and refined to a resolution of 1.9 Å (Table 1). The presence of Zinc ions in the crystals was confirmed by an X-ray absorption edge scan (Supplementary Figure S6).

The Spt5 NGN domain consists of a four-stranded anti-parallel β-sheet as part of an alpha–beta sandwich (Figure 1A). Spt4 also contains a four-stranded β-sheet stabilized by a zinc ion that is coordinated by the side chains of four conserved cysteine residues (Figure 1A). Mutations altering any of the four zinc-binding cysteines in Saccharomyces cerevisiae Spt4 cause loss-of-function phenotypes and in several instances greatly decrease cellular levels of Spt4 (38,39), suggesting that the zinc-binding site is an essential structural element of Spt4 (8). In the Spt4/5 NGN heterodimer, the β-sheets of Spt4 and Spt5 NGN are aligned edge to edge, forming a continuous eight-stranded anti-parallel β-sheet (Figure 1A). In addition to this β sheet interface, several hydrogen bonds contribute to the specificity of Spt4–Spt5NGN interactions. In particular, Ser39 from Spt4 forms a conserved hydrogen-bonding interaction with Glu46 in Spt5NGN (Figure 1B). This hydrogen bond fixes Glu46, which results in a strong acid–base interaction (Guo et al., 2008; Nicholson et al., 1988).

Additional Spt4-Spt5 interactions are formed by helix a1 and a2 of the Spt5 NGN domain (Figure 1D and Supplementary Figure S3). In particular, Ser39 from Spt4 forms a conserved hydrogen-bonding interaction with Glu46 in Spt5NGN (Figure 1B). This hydrogen bond fixes Glu46, which results in a strong acid–base interaction (Guo et al., 2008; Nicholson et al., 1988).

Additional Spt4-Spt5 interactions are formed by helix a1 and a2 of the Spt5 NGN domain (Figure 1D and Supplementary Figure S3). In particular, Ser39 from Spt4 forms a conserved hydrogen-bonding interaction with Glu46 in Spt5NGN (Figure 1B). This hydrogen bond fixes Glu46, which results in a strong acid–base interaction (Guo et al., 2008; Nicholson et al., 1988).
to a large extent between the *M. jannaschii* and *S. cerevisiae* heterodimers, including key conserved residues mediating the interaction (Figure 1B, C and Supplementary Figure S4).

To build a model of the complete archaeal Spt4/5, we have combined our structure of *M. jannaschii* Spt4/5 NGN and a homology model (‘Materials and Methods’ section) of the Spt5 KOW domain based on the structure of the *A. aeolicus* NusG KOW domain (Figure 1F). In the model, we maintained the relative orientations of the *M. jannaschii* Spt5 NGN and KOW domains observed in *A. aeolicus* NusG. The model of the *M. jannaschii* Spt5 KOW domain showed good stereochemistry with the parent—model Cα-RMSD of 0.7 Å. This is generally well correlated with model accuracy (41).

**Archaeal Spt4/5 stimulates transcription elongation in vitro**

We have recently developed single round transcription elongation assays for archaeal RNAP that make use of synthetic DNA–RNA elongation scaffolds (2). In order to test the effect of recombinant Spt4/5 on the transcription elongation properties of RNAP, we carried out transcription elongation assays using a recombinant 10-subunit core RNAP from *M. jannaschii* lacking RNAP subunits F/E, which are not required for promoter directed transcription (28). Like bacterial and eukaryotic RNAPs, the archaeal RNAP can utilize a combination of template DNA strand and RNA primer for transcription elongation, and the addition of the non-template strand (NTS) enhances the processivity (Figure 2A and B) (2). Spt4/5 stimulates the formation of the run-off transcripts both in the absence and in the presence of the non-template strand (Figure 2A and B). The stimulation is ~2.5-fold in the absence, and 1.75-fold in the presence of the NTS (Figure 2C and D), in good agreement with observations from the bacterial homologue NusG (17,42). The archaeal 12-subunit RNAP has a higher processivity compared to the 10-subunit enzyme, and congruent with this result its

![Figure 2](https://academic.oup.com/nar/article-abstract/38/12/4040/2409576)

**Figure 2.** Archaeal Spt4/5 stimulates transcription elongation. Transcription assay using a synthetic elongation scaffold and recombinant 10-subunit *M. jannaschii* RNAP (200 nM). Transcript synthesis was monitored over a time course of 2, 5, 10 and 20 min in the absence of NTS (A), and 20, 40, 90 and 300 s in the presence of NTS (B). Recombinant Spt4/5 (10 μM) was added to the reaction in conjunction with RNAP. The full length ‘run off’ transcript was quantitated in the absence (C) and presence of the NTS, and normalized to the reaction end point in the absence of Spt4/5 (20 min and 300 s w/o NTS, C, and plus NTS, D, respectively). The gels (A and B) are representative and all quantitations are based on at least three independent experiments (arbitrary units, AU).
response to Spt4/5 is near the detection limit of the assay [(2) and data not shown]. Our results demonstrate that archaeal Spt4/5 acts as a positive transcription elongation factor, and that its stimulation of elongation is not dependent on interactions with the NTS.

**Spt4/5 function requires the conserved NGN domain**

In order to characterize the functions of Spt4, and the Spt5 NGN and Spt5 KOW domains, we prepared domain deletion variants of the *M. jannaschii* Spt4/5 (‘Materials and Methods’ section). Based on structural and amino acid sequence alignments of bacterial, eukaryotic and archaeal NusG/Spt5 factors (Figure 1C), we split the *M. jannaschii* factor at the C-terminal end of helix 3 (at residue 82) into two domains Spt5 NGN (1–82) and Spt5 KOW (83–147), respectively. Recombinant forms of Spt5 NGN and Spt5 KOW were soluble and expressed at high levels, which suggests that the structural integrity of the individual domains is not severely compromised. We verified that the Spt5N variant is necessary and sufficient for stable complex formation with Spt4. Spt4 co-purifies with the Spt5 NGN domain, but not with Spt5 KOW, on affinity chromatography, and Spt4 and Spt5 NGN, but not Spt4 and Spt5 KOW, co-elute as a distinct complex from a size exclusion column (Supplementary Figure S1).

The heterodimeric interface between Spt4 and 5 has an area of 940 Å² and contains 10 hydrogen bonds and 4 salt bridges (Supplementary Figure S7). A mutagenesis analysis of residues at the interface showed that no single amino acid substitution was able to disrupt the complex (32). To test whether Spt4 had any influence on the stability of Spt5 at the reaction temperatures of the *M. jannaschii* system we tested their solubility following a 30 min incubation at 65 and 75°C. Whereas Spt4/5 and Spt4/5 NGN are soluble and heat stable at these temperatures, Spt5 and Spt5 NGN were soluble at 65°C but precipitated at 75°C (Figure 3). This shows that Spt4 stabilizes the Spt5 NGN domain, and offers an explanation of the temperature sensitive phenotype of *spt4* null mutants in yeast (43).

To identify the minimal domain configuration that is necessary and sufficient for stimulation of transcription elongation, we tested the Spt4/5 domain deletion variants in elongation assays. Three of the variants, Spt4/5, Spt4/5 NGN and Spt5 were able to stimulate transcription elongation, whereas Spt5 NGN, Spt5 KOW or Spt4 showed no activity (Figure 4 and data not shown). Our results show that the Spt5 NGN domain is strictly required, but not sufficient, for the stimulation of elongation. This result is reminiscent of the bacterial NusG and RfaH variants from *E. coli*. However, in contrast to the bacterial NGN domains, which are sufficient for the stimulation of transcription elongation (20), the archaeal Spt5 NGN domain requires either Spt4 or the Spt5 KOW domain for elongation activity.

**Figure 3.** Spt4 stabilizes the Spt5 NGN domain. The Spt4/5, Spt4/5 NGN, Spt5 and Spt5 NGN variants were incubated at either 65 or 75°C. The heat stable and soluble fraction was resolved on SDS-PAGE and the gel subsequently stained with Coomassie.

**Figure 4.** Domain deletion analysis identifies Spt5 NGN as effector domain. Transcription elongation assays using recombinant 10-subunit RNAP (200 nM) and Spt4/5, Spt4/5 NGN, Spt5, Spt5 KOW and Spt4 (each at 10 μM) in the absence of NTS. Samples were taken at 2, 5, 10 and 20 min (A and B). The full-length run-off transcript was quantitated and normalized to the end point (5 min) in the absence of any added factors (C).
Spt4/5 interacts with RNAP independently of nucleic acids

To investigate the interaction of Spt4/5 with RNAP, we established a semi-quantitative native gel electrophoresis assay using recombinant RNAP and radio-labelled Spt4/5 variants. Incubation of labelled Spt4/5 with increasing amounts of RNAP lead to a retardation of the Spt4/5 band and a new low mobility band indicating Spt4/5–RNAP complexes (Figure 5A). We tested the specificity of the Spt4/5–RNAP interaction by adding unlabelled Spt4/5 to the binding reaction, which resulted in an efficient competition of the Spt4/5–RNAP signal (Figure 5E). This demonstrates that the interaction of archaeal Spt4/5 with the RNAP, similar to NusG in bacteria does not depend on the nucleic acid scaffold of the elongation complex.

In order to identify the Spt4/5 domains that are involved in RNAP binding, we tested the wild-type and domain deletion variants of Spt4/5 in binding and competition experiments. Labelled Spt4/5 NGN, full-length Spt5 and Spt5 NGN were able to bind to RNAP, albeit with reduced affinity (Figure 5B–D). Spt4/5 and Spt4/5 NGN were able to compete for the binding of Spt4/5, whereas Spt5 KOW and Spt4 could not compete for the binding of Spt4/5 to RNAP (Figure 5E). We tested whether the RNAP subunits F/E contributed to the interaction with Spt4/5. Both the 10-subunit RNAP ΔF/E and the 12-subunit RNAP were equally able to recruit Spt4/5, implying that the interaction between Spt4/5 and RNAP is not substantially affected by F/E complex (Figure 5A).

In summary, our results suggest that archaeal Spt4/5 complex interacts with RNAP via the Spt5 NGN-domain. Since the fraction of Spt4/5 NGN that is shifted into RNAP-containing complexes is smaller than that of full-length Spt4/5, we suggest that the Spt5 KOW domain increases the affinity of Spt4/5 for RNAP. However, Spt5 KOW cannot alone compete for the binding of Spt4/5. Likewise, because the shift of Spt5 is weaker than that of Spt4/5, our results suggest that Spt4 increases the affinity for RNAP, but it cannot compete for the binding of the Spt4/5 to RNAP. Rather than directly contributing to RNAP binding, Spt4 could stabilize Spt5 NGN, which is congruent with our heat inactivation profile (Figure 3).

A hydrophobic depression on the NGN domain mediates binding to RNAP

The above data suggest that RNAP binding of Spt4/5 is mediated by the Spt5 NGN domain. Structures of the NGN domains of bacterial NusG and eukaryotic and archaeal Spt5 reveal a depression that is lined by hydrophobic amino acid side chains that have been proposed to interact with bacterial RNAP (17) (Figures 1, 6A and Supplementary FigureS5). We designed and introduced mutations into Spt5 that perturbed the hydrophobic nature of the putative RNAP binding site and tested the ability of the heterodimeric mutant Spt4/5 complexes to interact physically and functionally with the archaeal RNAP (Figure 6A–C). All recombinant Spt5 mutants were as soluble and heat stable as the wild-type Spt5 protein, which suggests that the structural integrity of Spt5 was not compromised by the amino acid substitutions. The Spt5 A4R mutation introduces both a bulky side chain and a positive charge into the hydrophobic cavity. The resulting Spt4/5-4 Arg variant cannot bind RNAP. Likewise, the Spt5 Y42A variant cannot bind RNAP. The Spt5 L44A mutation substitutes an isopropyl side chain with the smaller methyl group and renders Spt4/5-42 Ala incapable of binding RNAP. The Spt5 L44A mutation abrogates binding of Spt4/5-44 Arg to RNAP in our native gel assay. When we tested the Spt5 mutants in transcription elongation assays, the elongation phenotypes of the mutants showed the same trend as their RNAP binding properties (Figure 6C). Only the Spt4/5-44 Ala variant was able to stimulate elongation comparable to the wild-type Spt4/5 (110% of full-length transcript compared to wild-type Spt4/5). The activity of Spt4/5-42 Ala and Spt4/5-44 Arg was 65%, 60% and 80% of wild-type Spt4/5, respectively. Thus, all three mutants retained a stimulatory activity at levels above the negative control reactions (Figure 6C), which
Spt4/5 mutants (10 nM) Transcription elongation assays using 10-subunit RNAP (200 nM) and (50 and 400 nM) were separated with native gel electrophoresis (E. coli).

Figure 6. Site-directed mutagenesis of the Spt5 NGN domain identifies a hydrophobic cavity as RNAP binding site. The residues Ala-4, Tyr-42 and Leu-44 line a hydrophobic depression in the Spt5 NGN domain (A). Spt4/5-RNAP binding assays using labelled Spt4/5 mutants (Spt5-4Arg, -42 Ala, -44 Ala and -44 Arg, 50 nM) and 10-subunit RNAP (50 and 400 nM) were separated with native gel electrophoresis (B). Transcription elongation assays using 10-subunit RNAP (200 nM) and Spt4/5 mutants (10 μM) in the absence of NTS (C).

in all likelihood reflects the higher sensitivity of the elongation assay compared to the binding assay.

In summary, our site-directed mutagenesis analysis of the hydrophobic depression of the archaeal Spt5 NGN domain strongly suggests that it serves as a binding site for archaeal RNAP. Congruent with this result are mutations in the corresponding region of E. coli NusG, which also interfere with stimulatory effect of NusG on transcription elongation (17) even though they have not been shown to interfere with the binding of NusG to RNAP directly. Thus, a conserved hydrophobic depression on the Spt5 NGN domain surface is required for RNAP binding and elongation stimulation.

Spt4/5 interacts with archael RNAP via the coiled-coil motif of the RNAP clamp

Escherichia coli NusG and its parologue RfaH interact with the bacterial RNAP via a coiled-coil motif of its flexible RNAP clamp (E. coli RNAP β’-residues 265–310) (21,44). In order to test whether the same site on archael RNAP is used for Spt4/5 binding, we replaced the ten amino acids ( mjA’ 255–264) corresponding to the tip of the coiled coil with a tetra-glycine linker, and tested the resulting RNAPCC-Gly4 variant for binding and elongation stimulation by Spt4/5 (Figure 7). The catalytic activity of the RNAP CC-Gly4 variant is identical to the wild-type enzyme in promoter independent transcription assays (data not shown) and both quantitatively and qualitatively indistinguishable in transcription elongation assays (Figure 7A), which demonstrates that the reconstitution of the recombinant RNAP and the structural integrity of the active site is not compromised by the CC-Gly4 mutation, like the corresponding mutation in bacterial RNAP (21). Whereas transcription elongation of the wild-type RNAP is stimulated by Spt4/5, RNAP CC-Gly4 is unable to respond to Spt4/5 (Figure 7A–D) and unable to recruit Spt4/5 in binding assays (Figure 7E). In summary, both functional and physical interactions between Spt4/5 and RNAP are abrogated by a mutation in the tip of RNAP clamp coiled-coil motif at the rim of the active centre cleft, suggesting this site as a conserved interaction site of RNAPs.

DISCUSSION

Like the core RNAP subunits, Spt5 is present in all three kingdoms of life and is one of the most ancient components of the transcription machinery (8,45,46). In particular, the NGN domain of Spt5 shows striking structural similarity between bacteria, archaea and eukaryotes. We show here that this similarity extends to the mechanisms that Spt5 uses to bind RNAP and to promote transcription. We provide evidence that a conserved interaction between the tip of the RNAP clamp coiled-coil and a hydrophobic depression on the Spt5 NGN domain mediates complex formation and elongation stimulation.

We have identified three hydrophobic residues (alanine-4, tyrosine-42 and leucine-44) in the archael Spt5 NGN domain that are apparently part of the RNAP-binding site. These hydrophobic residues line a concave depression that is located opposite of the Spt4 interaction surface and oriented away from the Spt5 KOW domain of our model of full-length Spt4/5, consistent with RNAP binding without steric hindrance. We also showed that replacement of ten amino acids forming the tip of the clamp coiled-coil motif of the largest subunit of the M. jannaschii RNAP (A’ residues 255-264) with a tetra glycine linker abolishes Spt4/5 binding and Spt4/5 elongation stimulation. Consistently, a similar mutation interferes with the response of bacterial RNAP to the NusG paralogue RfaH (21). Also consistent, a coiled-coil fragment of the E. coli RNAP (β’ 262–309) interacts with NusG (residues 1–132) in a genetic assay (44), and the NusG NGN domain mediates RNAP binding and elongation stimulation (17). The elongation-stimulatory activity of Spt4/5 and NusG are generally mild in vitro. Whereas we measured a 1.8-fold increase in the formation of the runoff transcript, Landick and colleagues (17) report a 1.2-fold increase, and Nudler and colleagues (42) obtain a 1.8-fold increase in the elongation rate of bacterial RNAP.

The elongation–stimulatory mechanism of Spt4/5 is likely allosteric, since the Spt4/5 binding site on the tip of the RNAP clamp coiled-coil is located far above
active centre on one rim of the polymerase cleft. Since this site is close to the path of the NTS in the elongation complex (47), Spt4/5 could modulate elongation by altering RNAP–nucleic acid interactions, maybe by facilitating DNA–RNA strand separation or upstream DNA strand rewinding. However, our observation that Spt4/5 stimulates elongation in the absence of the NTS argues against these mechanisms. Alternatively, Spt4/5 could influence the active centre by altering the clamp position, as suggested for the bacterial NusG paralogue RfaH (21,48). Details of the Spt4/5 mechanism remain to be investigated, including the mechanism of transmitting

Figure 7. A deletion-substitution (CC-Gly4) in the tip of the RNAP clamp coiled-coil domain abrogates Spt4/5 binding and stimulation. Transcription elongation assays with 10-subunit RNAP and RNAP CC-Gly4 in the absence (A, 200 nM) and presence of NTS (C, 50 nM), ± Spt4/5 (10 μM). Run-off transcripts were quantitated and normalized to the end points of reactions carried out without Spt4/5 (B and D). The binding of Spt4/5 to RNAP and RNAP CC-Gly4 was monitored with native gel electrophoresis (E). The location of the RNAP clamp coiled-coil (highlighted in red), the active site NADFDGD motif (green) and metal ion (magenta) in archaeal RNAP (Sulfolobus shibatae RNAP, pdb 2WAQ).
an allosteric signal, and the question whether Spt4/5 stimulates elongation at the level of transcription rate, RNAP processivity, or both. In eukaryotic transcription systems it is exceedingly difficult to directly monitor the effect of Spt4/5 on elongation by RNAPII, which require additional factors including the Paf-1 complex and Tat-SF1 (49). Thus, the archaerial transcription systems offer a unique opportunity to unravel the molecular mechanisms of Spt4/5.

Comparison of our data with published results reveals interesting differences between archaerial and eukaryotic Spt4/5. In particular, archaerial Spt4/5 binds directly to RNAP and independently of the DNA–RNA nucleic acid scaffold, whereas human Spt4/5 (DSIF for DRB-sensitivity inducing factor) binding to RNAPII apparently depends on the presence of the RNA transcript (50), although another study showed that association of Spt4/5 with elongating RNAPII was insensitive to nuclease treatment (51). An RNA-dependency in eukaryotic systems could be due to interactions between the transcript and Spt5 domains that are not conserved in the archaerial variant, including the additional copies of KOW domain and C-terminal repeats. Alternatively, the interactions between the hyperthermophilic archaerial Spt4/5 and RNAP may be stronger owing to more stringent requirements for the stability of RNAP–transcription factor complexes at the biologically relevant high reaction temperatures of *M. jannaschii*. Similarly, the interactions between RNAP subunits in *M. jannaschii* are stronger when compared to *S. cerevisiae* RNAPII subunits (23).

Whereas the N-terminal NGN domain of NusG/Spt5 mediates a conserved RNAP contact, the C-terminal KOW domain(s) in these factors apparently mediate additional, species-specific interactions. In the bacterial system, an activity of NusG during transcription termination is dependent on both NGN and KOW domains, but especially the latter has been implicated in facilitating the recruitment and/or interactions with accessory factors that include Rho, NusA, NusB and NusE (17, 22). Eukaryotic Spt5 harbours multiple copies of the KOW domain that may mediate some of the multiple physical and functional interactions of Spt4/5, including interactions with the negative transcription elongation factor NELF, Spt5 and RNAPII by phosphorylation, including P-TEFb and Bur-1/Bur-2, and factors that are involved in chromatin remodelling such as Spt6 and FACT, and RNA processing factors such as the mRNA capping enzyme and the cap methyl transferase (51). Hitherto no factors have been identified that interact with the Spt5 KOW domain in the Archaea.

Our results reveal that the structure and the elongation function of Spt4/5 and NusG, including the interaction site on the distinct RNAPs, are universally conserved throughout the three domains of life. Considering that eukaryotic Spt4/5 regulates both RNAPII and RNAPII (52), Spt4/5 seems evolutionary ancient, and the molecular mechanisms by which it regulates gene expression are likely to have their origin prior to the diversification of distinct classes of eukaryotic RNA polymerases, and indeed before the split of the three main lineages of life, Bacteria, Archaea and Eukarya. Since the transcription factors that govern transcription initiation in Bacteria and Archaea/Eukarya are not homologous, this lends itself to the hypothesis that the regulation of transcription elongation might even precede the regulation of transcription initiation.

### ACCESSION NUMBER

Our entry ‘Crystal structure of Spt4/5NGN heterodimer complex from *Methanococcus Jannaschii*’ has been assigned the RCSB ID code rcsb057564 and PDB ID code 3LPE.

### SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

### ACKNOWLEDGEMENTS

We would like to thank Dr Kristine Bourke Arnvig for critical reading of the article. We thank Clemens Schulze-Briese and the SLS for assistance. Part of this work was performed at the Swiss Light Source (SLS) at the Paul-Scherrer-Institut in Villigen, Switzerland. We also thank Elena Conti and her group for use of their crystallization facility (Max Planck Institute for Biochemistry, Martinsried, Germany).

### FUNDING

The work at UCL ISMB was funded by Wellcome Trust project grant 079351/Z/06/Z and BBSRC project grant BB/E008232/1 awarded to F.W. D.K. would like to thank Heinz-Jürgen Steinhoff at the University of Ottnadbrück for financial support by DFG Sonderforschungsbereich 431 P18. G.E.D. was supported by the Elite Netzwerk Bayern. Funding for open access charge: Wellcome Trust and University College London.

Conflict of interest statement: None declared.

### REFERENCES

1. Werner,F. (2008) Structural evolution of multisubunit RNA polymerases. *Trends Microbiol.*, 16, 247–250.
2. Hirtreiter,A., Grohmann,D. and Werner,F. (2009) Molecular mechanisms of RNA polymerase II (RPB4/7) complex is required for high processivity in vitro. *Nucleic Acids Res.*, 38, 585–596.
3. Landick,R. (2006) The regulatory roles and mechanism of transcriptional pausing. *Biochem. Soc. Trans.*, 34, 1062–1066.
4. Hartzog,G.A., Speer,J.L. and Lindstrom,D.L. (2002) Transcript elongation on a nucleoprotein template. *Biochim. Biophys. Acta*, 1577, 276–286.
5. Hartzog,G.A. (2003) Transcription elongation by RNA polymerase II. *Curr. Opin. Genet. Dev.*, 13, 119–126.
6. Borukhov,S., Lee,J. and Laptenko,O. (2005) Bacterial transcription elongation factors: new insights into molecular mechanism of action. *Mol. Microbiol.*, 55, 1315–1324.
7. Werner,F. (2007) Structure and function of archaerial RNA polymerases. *Mol. Microbiol.*, 65, 1393–1404.
8. Guo,M., Xu,F., Yamada,J., Egelhofer,T., Guo,Y., Hartzog,G.A., Teng,M. and Niu,L. (2008) Core structure of the yeast spt4-spt5
complex: a conserved module for regulation of transcription elongation. **Structure, 16**, 1649–1658.

9. 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.

10. Ivanov,D., Kwak,Y.T., Guo.J. and Gaynor,R.B. (2000) Domains in the SPT5 protein that modulate its transcriptional regulatory properties. **Mol. Cell Biol., 20**, 2970–2982.

11. Hartzog,G.A., Wada,T., Handa,H. and Winston,F. (1998) Evidence that Sp4, Sp5, and Sp6 control transcription elongation by RNA polymerase II in Saccharomyces cerevisiae. **Genes Dev., 12**, 357–369.

12. Wada,T., Takagi,T., Yamaguchi,Y., Ferdous,A., Imai,T., Hirose,S., Sugimoto,S., Yano,K., Hartzog,G.A., Winston,F. et al. (1998) DSIF, a novel transcription elongation factor that regulates RNA polymerase II processivity, is composed of human Sp4 and Sp5 homologs. **Genes Dev., 12**, 343–356.

13. Core,L.J. and Lis,J.T. (2008) Transcription regulation through promoter-proximal pausing of RNA polymerase II. **Science, 319**, 1791–1792.

14. Margaritas,T. and Holstege,F.C. (2008) Pooled RNA polymerase II gives pause for thought. **Cell, 133**, 581–584.

15. Artsimovich,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.

16. Burov,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.

17. Steiner,T., Kaiser,J.T., Marinkovic,S., Huber,R. and Wahl,M.C. (2009) The RNAP elongation factors. **Mol. Cell Biol., 29**, 4641–4653.

18. Vagin,A.A., Steiner,R.A., Lebedev,A.A., Potterton,L., McNicholas,S., Long,F. and Murshudov,G.N. (2004) REFMAC5 dictionary: organization of prior chemical knowledge and guidelines for its use. **Acta Cryst. D Biol. Crystallogr., 60**, 2184–2195.

19. Werner,F. and Weis,G. (2005) Direct modulation of RNA polymerase core functions by basal transcription factors. **Mol. Cell Biol., 25**, 8344–8355.
51. Lindstrom, D.L., Squazzo, S.L., Muster, N., Burckin, T.A., Wachter, K.C., Emigh, C.A., McCleery, J.A., Yates, J.R. III and Hartzog, G.A. (2003) Dual roles for Spt5 in pre-mRNA processing and transcription elongation revealed by identification of Spt5-associated proteins. Mol. Cell Biol., 23, 1368–1378.

52. Schneider, D.A., French, S.L., Osheim, Y.N., Bailey, A.O., Yu, 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.