

# Structural basis of archaeal RNA polymerase transcription elongation and Spt4/5 recruitment

Daniela Tarău<sup>1</sup>, Felix Grünberger<sup>1,†</sup>, Michael Pils<sup>2,†</sup>, Robert Reichelt<sup>1,†</sup>, Florian Heiβ<sup>2</sup>, Sabine König<sup>3</sup>, Henning Urlaub<sup>3,4</sup>, Winfried Hausner<sup>5,1</sup>, Christoph Engel<sup>2,\*</sup> and Dina Grohmann<sup>1,5,\*</sup>

<sup>1</sup>Institute of Microbiology & Archaea Centre, Single-Molecule Biochemistry Lab, University of Regensburg, 93053 Regensburg, Germany

<sup>2</sup>Regensburg Center for Biochemistry (RCB), Structural Biochemistry Group, University of Regensburg, Regensburg, Germany

<sup>3</sup>Bioanalytic Mass Spectrometry, Max Planck Institute for Multidisciplinary Sciences, 37077 Göttingen, Germany

<sup>4</sup>Department of Clinical Chemistry, University Medical Center Göttingen, 37075 Göttingen, Germany

<sup>5</sup>Regensburg Center for Biochemistry (RCB), University of Regensburg, 93053 Regensburg, Germany

\*To whom correspondence should be addressed. Tel: +49 941 943 3147; Fax: +49 941 943 2403; Email: dina.grohmann@ur.de

Correspondence may also be addressed to Christoph Engel. Tel: +49 941 943 2718; Fax: +49 941 943 2474; Email: christoph.engel@ur.de

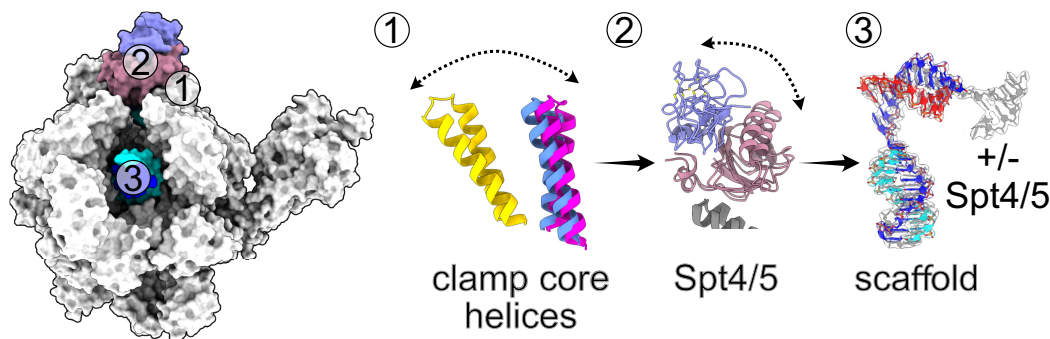
†The second, third and fourth author should be regarded as Joint Second Authors.

## Abstract

Archaeal transcription is carried out by a multi-subunit RNA polymerase (RNAP) that is highly homologous in structure and function to eukaryotic RNAP II. Among the set of basal transcription factors, only Spt5 is found in all domains of life, but Spt5 has been shaped during evolution, which is also reflected in the heterodimerization of Spt5 with Spt4 in Archaea and Eukaryotes. To unravel the mechanistic basis of Spt4/5 function in Archaea, we performed structure-function analyses using the archaeal transcriptional machinery of *Pyrococcus furiosus* (*Pfu*). We report single-particle cryo-electron microscopy reconstructions of apo RNAP and the archaeal elongation complex (EC) in the absence and presence of Spt4/5. Surprisingly, *Pfu* Spt4/5 also binds the RNAP in the absence of nucleic acids in a distinct super-contracted conformation. We show that the RNAP clamp/stalk module exhibits conformational flexibility in the apo state of RNAP and that the enzyme contracts upon EC formation or Spt4/5 engagement. We furthermore identified a contact of the Spt5-NGN domain with the DNA duplex that stabilizes the upstream boundary of the transcription bubble and impacts Spt4/5 activity *in vitro*. This study, therefore, provides the structural basis for Spt4/5 function in archaeal transcription and reveals a potential role beyond the well-described support of elongation.

## Graphical abstract

### Structural characterisation of *Pfu* RNAP complexes



## Introduction

During the first step in gene expression, genomic DNA is transcribed in a highly regulated manner to synthesize the cellular RNA pool. The cyclic and multistep transcription process can be divided into three main steps: initiation, elongation, and termination (1–5). Although Archaea are prokaryotes, the components of the archaeal transcription machin-

ery do not resemble the bacterial transcription apparatus but are highly homologous to the eukaryotic RNAP II transcription machinery (3,4). This applies to the composition and structure of the RNAP itself and extends to the basal transcription initiation and elongation factors. Eukaryotic Pol I and III even retain structural elements lost in Pol II, suggesting a common archaeal origin (6–8). In Archaea, tran-

Received: January 2, 2024. Revised: March 25, 2024. Editorial Decision: March 27, 2024. Accepted: April 4, 2024

© The Author(s) 2024. Published by Oxford University Press on behalf of Nucleic Acids Research.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License

(https://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

scription initiation depends on the TATA-binding protein (TBP) and transcription factor B (TFB), which recognize the promoter DNA, forming sequence-specific interactions with the TATA-box and B-recognition element (BRE), respectively (9–15). The helicase-independent melting of the promoter DNA around the transcription start site is supported by transcription factor E (TFE), leading to the efficient formation of the open pre-initiation complex (PIC) (10,16–19). Productive RNA synthesis becomes possible when contacts between initiation factors and RNAP are broken, and RNAP escapes from the promoter.

Transition into the elongation phase is accompanied by the replacement of TFE by the elongation factor Spt4/5 (11,12,17). Spt5 is the only conserved transcription factor found in all domains of life (3). Like its bacterial counterpart NusG, archaeal Spt5 comprises an N-terminal NusG domain (NGN), a linker domain, and a C-terminal Kypridis-Ouzounis-Woese (KOW) domain (20). Archaeal and eukaryotic Spt5 form stable heterodimers with the small protein Spt4 (20–23). In Eukaryotes, however, Spt5 is much larger and includes four to seven consecutive KOW domains at its C-terminus. Productive RNA synthesis during the elongation phase is not continuous but is frequently interrupted by pausing events. Upon interaction with the elongating polymerase, Spt4/5 reduces pausing, thereby increasing the processivity of RNAP (24). Moreover, Spt4/5 promotes transcription through chromatinized DNA templates in Archaea and Eukaryotes (20,25). In Bacteria, NusG suppresses pausing and backtracking of the RNAP and interacts with effector proteins such as the termination factor Rho via its KOW domain (26,27). Furthermore, it was shown that NusG mediates the coupling between RNAP and the ribosome (28–31). In contrast to NusG in *E. coli*, gene-specific homologues called RfaH and ‘NusG-like’ found in *B. subtilis* interact with the non-template strand of the transcription bubble and thereby enhance transcriptional pausing (32–35).

Structural, single-molecule and biochemical data revealed the mechanistic basis for the stimulatory effect of NusG and Spt4/5. Spt5, as well as NusG, binds to the tip of the RNAP clamp coiled-coil motif via a hydrophobic cavity in the NGN domain and bridges over the DNA binding channel to the RNAP protrusion domain (21,22). The clamp is a domain mostly conserved in structure among multi-subunit RNAPs and can adopt open or closed conformations. While the degree of clamp closure varies among transcription systems, especially by the eukaryotic RNAPs, a closed clamp defines actively elongating RNAPs. At the same time, an expanded or partially expanded conformation is found in inactive RNAPs (36,37). NusG prevents conformational changes that are connected to pausing states, e.g. the opening of the clamp (35). Archaeal Spt4/5 induces the closure of the clamp and contraction of the RNAP when the initiation complex transitions to the elongation complex, thereby securing the DNA template in the DNA binding channel and rendering the elongation complex more stable (19). Furthermore, Spt4/5 and RNAP subunits Rpo4/7 are required for efficient termination mediated by the extrinsic termination factor aCPSF1 in Archaea, suggesting that both protein complexes provide interaction sites for aCPSF1 (38,39).

High-resolution crystal structures and structural models of bacterial and eukaryotic transcription complexes derived from cryo-electron microscopy (cryo-EM) and single particle analysis data provided insights into the molecular basis

of transcription initiation, elongation, transcriptional pausing, and termination (1,5,40–45). However, structural information on archaeal transcription complexes is sparse and only includes structures of the RNAP apo enzyme (46–50), of RNAP in complex with DNA (51) or TFE, as well as a complex formed of RNAP, DNA and TFE (49). Furthermore, structures of RNAP in complex with TFS4 (52) or an archaeoviral inhibitor (50) were published. Additionally, an isolated clamp fragment in complex with archaeal Spt4/5 (21) was structurally described, and a low-resolution structural model of the archaeal RNAP-Spt4/5 complex in the absence of nucleic acids is available, but structural details could not be discerned (22). Distance constraints derived from a comprehensive single-molecule FRET study were used to build a structural model of the full archaeal PIC (16).

To gain insights into the structural (re-)organization of archaeal transcription elongation complexes (TEC) in the absence and presence of Spt4/5 and to understand the mechanisms of Spt4/5 action, we determined single-particle cryo-EM reconstructions of the archaeal RNAP apo enzyme from *P. furiosus* with and without Spt4/5 and the archaeal TEC with and without Spt4/5. Combined with biochemical and mass spectrometry data, we found that a subfraction of the cellular apo RNAP pool forms a complex with Spt4/5 in the absence of DNA or RNA. In this complex, the RNAP adopts a supercontracted conformation not described for other RNAPs. In the context of the TEC, Spt4/5 occupies a slightly different position relative to the RNAP compared to the apo RNAP-Spt4/5 complex. Furthermore, we found that archaeal Spt4/5 interacts with the upstream DNA, representing an organism-specific interaction mode. Transcription assays with Spt5 mutants revealed that this interaction contributes to the processivity of the archaeal RNAP. Our data, therefore, add to the structural and mechanistic understanding of archaeal transcription complexes and the functional role of Spt5.

## Materials and methods

### Synthetic oligonucleotides

Elongation complexes were formed using an elongation scaffold composed of a fully complementary template and non-template strand and a short (14 nt) or slightly longer (22 nt) RNA oligonucleotide complementary to the 9 bases (CGCCTGGTC) of the template strand. All DNA and RNA oligonucleotides were synthesized by Microsynth AG. Sequences are as follows: template strand (TS) 5'-CCACCCTTACCTCCACCATATGGGAGATCCATTAGAGTAGCGCCTGGTCATTACTTAAAGATGAA GTAGTTATAGTACTGCCGG-3'; non-template strand (NTS) 5'-CCGGCAGTACTATAACTACTTCATCTTAAGTAATGACCAGGCGTACTCTAATGGATCTCCCATATGGTGGAGGTAAGGGTGG-3'; partially complementary RNA (short) 5'-AUUUAGACCAGGCG-3'; and long RNA 5'-UUUUUUUUUUUAGACCAGGCG-3'.

### Genomic modification of the *rpo3* gene in *P. furiosus* and cultivation of *Pyrococcus furiosus*

To enable a simplified purification of the cellular RNAP, we introduced a StrepII-His6-tag at the C-terminal end of subunit Rpo3 using an established genetic system for *P. furiosus* strain DSM3638 (53,54). We used double crossover recombination with plasmid pMUR1 to modify the genome. This plasmid

contains *rpo3* (PFDSM3638\_08330) together with 250 bp of the upstream region. The corresponding sequence was amplified from wild-type DNA with the primer pair 1648/1647-BH1-F and PF1647\_Fus-StrepII\_Ri (Supplementary Table S1). The StrepII-His6-tag at the C-terminal end was provided by a double-stranded oligonucleotide (primer Fus\_Strep\_His\_F and Fus\_Strep\_His\_R). For the selection of the *Pyrococcus* transformants with the antibiotic simvastatin, the coding sequence of the hydroxymethylglutaryl CoA reductase from *Thermococcus kodakarensis* was used. The corresponding sequence was amplified from plasmid pMUR27 using the primer pair PF1647upTk0914F and TK0914-R (54). Furthermore, the plasmid includes a 950 bp sequence downstream of *rpo3*, enabling double crossover recombination. This fragment was amplified from genomic DNA using the primer pair Tk0914PF1647dwF and Fus004\_BamHI-R. All DNA fragments were combined step-by-step by single overlap extension PCR. The fused fragment was hydrolyzed with BamHI and ligated into a pUC19 vector, which was also hydrolyzed with BamHI and dephosphorylated with calf intestinal phosphatase.

The resulting plasmid pMUR1 was sequenced, and 1  $\mu$ g of the linearized version was used for the transformation of *P. furiosus* as described previously (53,54). Transformants were enriched with 10  $\mu$ M simvastatin in  $\frac{1}{2}$  SME-starch liquid medium at 85°C for 48 h, and pure cultures were isolated after plating the cells on solidified medium in the presence of 10  $\mu$ M simvastatin. The integration of the DNA fragment into the genome via double crossover was verified by sequencing of the corresponding PCR products.

*P. furiosus* cells were first grown for 6 h in a volume of 20 ml under anaerobic conditions in  $\frac{1}{2}$  SME medium supplemented with 0.1% peptone, 0.1% yeast extract and 0.1% starch according to previously established protocols (55). The pre-culture was incubated at 90°C with a N<sub>2</sub> gas phase at 200 kPa pressure until a cell density of  $7.5 \times 10^6$  cells was reached. The pre-culture was used to inoculate a 100 L fermenter with  $\frac{1}{2}$  SME medium supplemented with 0.1% peptone, 0.1% yeast extract and 0.1% starch. The pH was adjusted with 5 M NaOH to 7.5, and cells were incubated for 12 h at 85°C and stirred at 150 rpm. The gas phase was N<sub>2</sub> (200 kPa). 117 grams of pellet were obtained after harvesting by centrifugation at 20 144 g, with Centrifuge 17S/RS (Heraeus Sepatech) with a continuous flow rotor 3049 (titanium). The pellet was split into three aliquots, plunge-frozen and stored at -80°C until further use.

### Purification of RNA polymerase from *Pyrococcus furiosus*

The cell pellet was resuspended in 2.5 ml/g of cells lysis buffer (100 mM Tris/HCl pH 8.0, 1 M NaCl, 20 mM imidazole, 20% (v/v) glycerol, 2.5 mM MgCl<sub>2</sub>) and 1 EDTA free protease inhibitor cocktail tablet (Roche) per 50 ml of lysis buffer. 0.5 g of glass beads (0.106  $\mu$ m diameter, Sigma-Aldrich) were added to the mixture, and the pellet was resuspended by bead beating using the FastPrep-24™ 5G bead beating grinder and lysis system from MP at a speed of 6.5 m/sec for 40 s. Cells were lysed by sonication using Bandelin Sonopuls HD 2070 at 60% amplitude for 6 min. Sonication was repeated five to six times. Cell lysis was confirmed by light microscopy. After lysis, the crude extract was centrifuged

at 15 000 g for 10 min (Eppendorf 5810 centrifuge) to pellet cell debris and glass beads. The supernatant was transferred and subjected to ultracentrifugation at 100 000 g using a type 70 Ti Rotor in the Optima Le-80K Ultracentrifuge (Beckmann).

The supernatant was carefully collected and filtered using 0.45  $\mu$ m syringe filters (Cytiva). The RNAP was purified using affinity chromatography exploiting the 6xHis-tag. To this end, the filtrate was applied to a HisTrap™ 5 ml Fast Flow column (Cytiva) equilibrated in lysis buffer using an Äkta Purifier system (GE Healthcare). The column was washed with lysis buffer (100 mM Tris/HCl pH 8.0, 1 M NaCl, 20 mM imidazole, 20% (v/v) glycerol, 2.5 mM MgCl<sub>2</sub>) until the absorbance at 280 nm dropped close to 0 mAU. The protein was eluted from the column using 30 ml elution buffer (100 mM Tris/HCl pH 8.0, 300 mM NaCl, 250 mM imidazole, 20% (v/v) glycerol, 2.5 mM MgCl<sub>2</sub>). Subsequently, the RNAP was further purified, performing a size exclusion chromatography step (column: HiLoad 26/600 Superdex 200 PG from Sigma-Aldrich). The following running buffer was used for the size exclusion chromatography: 100 mM HEPES/NaOH pH 8.0, 150 mM NaCl, 20% (v/v) glycerol, 2.5 mM MgCl<sub>2</sub>. Elution fractions were loaded on a denaturing SDS polyacrylamide gel (PAGE) (Supplementary Figure S1), and fractions containing the RNAP were aliquoted and stored at -80°C until further use. The concentration of 1.638 mg/ml was determined by absorption spectroscopy at a wavelength of 280 nm using a NanoDrop™ One/OneC Microvolume UV-Vis Spectrophotometer (Thermo Fisher Scientific).

### Recombinant Spt4/5 expression and purification

The genes encoding Spt4 (PFDSM3638\_01210) and Spt5 (PFDSM3638\_10025) were amplified by PCR from genomic DNA of *P. furiosus*. The PCR fragment was cloned into pET14b and pET17b expression vectors, respectively by Gibson Assembly (56). The *spt4* gene was genetically fused to a sequence that encodes a 6xHis-tag at the C-terminus of Spt4, whereas Spt5 was cloned without any additional tag. 200 ng of each plasmid were used to transform *Escherichia coli* Rosetta™ (DE3) competent cells (Merk, Darmstadt) using heat transformation. The cell suspension was plated on LB-agar plates containing 100  $\mu$ g/ml ampicillin. Plates were incubated at 37°C overnight. The next day, one positive colony from each plate was resuspended in 20  $\mu$ l LB medium and plated again on an LB agar plate that contained 100  $\mu$ g/ml ampicillin and the plate was incubated again at 37°C overnight. The next day, the bacterial colonies were resuspended in 5 ml LB liquid medium, and this solution was used to inoculate a 1 l culture in LB medium using a starting OD<sub>600</sub> of ~0.07. The cultures were incubated at 37°C shaking at 160 rpm. Protein expression was induced by the addition of isopropyl-beta-D-thiogalactosid (IPTG) to a final concentration of 1  $\mu$ g/ml at an OD<sub>600</sub> of 0.6–0.7. After induction, the cells were further incubated for 4 h at 37°C shaking at 160 rpm. Cells were harvested by centrifugation for 10 min at 5500 g at 4°C. The pellets were resuspended in low salt buffer (40 mM HEPES/KOH pH 7.4, 50 mM NaCl, 15% (v/v) glycerol) and sonicated five times for 30 sec using a Bandelin Sonopuls HD 2070 at 50% amplitude and cycle 5. Cell debris was eliminated by centrifugation (1 h; 20 000 g; 4°C). Spt4 and Spt5 maintain their ther-

mostability even when recombinantly expressed. Therefore, the crude extract was incubated at 80°C for 20 min to heat denature and precipitate most of the *E. coli* proteins. The precipitates were pelleted by centrifugation (20 000 g; 15 min; 4°C) and the supernatant was filtered using a 0.2 µm syringe filter (Cytiva). At this point, the lysates that contained either Spt4 or Spt5 were mixed. The Spt4/5 complex was purified via affinity chromatography employing the 6xHis-tag of Spt4. Affinity purification was performed using the Äkta Purifier system and a HisTrap™ 1 ml Fast Flow column (Cytiva) equilibrated in low salt buffer. The lysate was loaded onto the column and washed in low salt buffer until the absorbance at 280 nm dropped close to 0 mAU. Spt4/5 was eluted using a gradient elution (elution buffer: 40 mM HEPES/KOH pH 7.4, 100 mM NaCl, 15% (v/v) glycerol, 500 mM imidazole). Elution fractions containing Spt4/5 were identified using SDS PAGE analysis and subjected to size exclusion chromatography (Superdex® 75 10/300 GL; Sigma-Aldrich). The size exclusion run was performed using isocratic elution in low salt buffer (40 mM HEPES/KOH, pH 7.4, 50 mM NaCl, 15% glycerol (v/v)). The purification success was assessed by SDS-PAGE. The concentration was determined using absorption spectroscopy, and an extinction coefficient of 15 930 was calculated with ProtParam tool (Expasy) (57).

#### Spt5 mutagenesis

Residues H12, R64, H65 and R67 were mutated to alanine via mutagenic primer PCR. The forward mutagenesis primer was phosphorylated at the 5' end and contained the mutated sequence at the respective positions. The reverse primer was designed to cover the complementary sequence just upstream of the forward primer. These primers were used in a PCR reaction to amplify the full plasmid. After gradient PCR, the successful amplification was verified by 1% agarose gel electrophoresis. 5 µl from the PCR product was circularized using 1 µl T4 DNA ligase and 2 µl ligation buffer 10× (NEB) in a final volume of 20 µl. To stop the ligation reaction, the sample was incubated at 60°C for 1 h to denature the T4 DNA ligase. 5 µl from the ligation mixture was used to transform 100 µl *E. coli* DH5α competent cells. Transformants were selected on 100 µg/ml ampicillin-containing LB agar plates (incubation of plates overnight at 37°C). A colony was collected and resuspended in 5 ml liquid LB medium with 100 µg/ml ampicillin and cultivated overnight at 37°C. Plasmid isolation was performed using the Plasmid Miniprep Kit, peqGOLD Plasmid Miniprep Kit I (VWR Life Science). Successful mutagenesis was verified by Sanger sequencing.

#### Spt5 mutant purification

Spt5 mutants were overexpressed as described for the wild-type protein. Heterodimer formation with Spt4 was performed on a HisTrap™ 1 ml Fast Flow (Cytiva) to which 1 mg of purified Spt4-6xHis protein was pre-loaded and immobilized. Bacterial lysate containing a Spt5 mutant was washed over the column. The Spt4/5 complex was eluted using a gradient elution (elution buffer consisted of 40 mM HEPES/KOH pH 7.4, 100 mM NaCl, 15% v/v glycerol, 500 mM imidazole). The Spt4/5 heterodimer was further purified by size exclusion chromatography (Superdex 75 10/300 GL, GE Healthcare) through isocratic elution in low salt buffer. Fractions containing Spt4/5 were evaluated and identified by SDS-PAGE.

#### Elongation complex assembly for cryo-electron microscopy analysis

Transcription elongation complexes (TECs) were assembled by first annealing the RNA (400 µM) to the TS (400 µM) using a molar ratio of 9:1 in a final volume of 40 µl. This sample was incubated at 95°C for 3 min and slowly cooled down. In a second step, RNAP was added to the pre-annealed RNA-TS hybrid (40 µM) using a 1:3 molar ratio by incubation at 80°C for 15 min. In a last step, the NTS is added to the complex with 3 times excess in molarity with respect to the RNAP, and this sample was incubated at 80°C for 10 min. The final volume of the reaction was 100 µl. For the assembly of the TEC in the absence of Spt4/5, a short RNA was used for the elongation scaffold, whereas for the TEC with Spt4/5 sample, a long RNA oligonucleotide was used instead. To increase the occupancy of Spt4/5 in the TEC, a BS<sup>3</sup> crosslinker (Thermo Fisher Scientific) was added to the preformed complex at a final concentration of 3 mM. TEC and Spt4/5 were incubated at 30°C for 30 min. The reaction was quenched by incubation at 30°C for 30 min with Lys/Asp mixture (0.1 mM) and incubation at 20°C for 20 min with ammonium bicarbonate (0.02 mM). The excess of crosslinker, quencher and precipitates were eliminated by gel filtration chromatography using the same column and buffer as described for the purification of Spt4/5 wt. The crosslinking efficiency was checked using a 4–20% gradient SDS-PAGE.

#### Cryo-EM grid preparation, data acquisition and processing

The overall workflow described in Pilsel *et al.* (58) was adapted to *Pfu* RNA polymerase complexes.

#### Apo RNA polymerase

100 µg of apo RNAP were run on a Superose 6 Increase 3.2/300 column in cryo-buffer (100 mM HEPES/KOH pH 8.0, 100 mM NaCl, 2.5 mM MgCl<sub>2</sub>). The 50 µl elution fractions corresponding to the polymerase peak were collected and used for vitrification (Supplementary Figure S2a, b). The sample was applied onto freshly glow discharged with EasiGlow, (TedPella) (1 × 0.4 mbar; 15 mA; 30 s) UltraAufoil R1.2/1.3, 300 mesh gold grids (Quantifoil) and plunge-frozen into liquid ethane (Vitrobot Mark IV; Thermo Fischer Scientific; 100% humidity; 4°C; 3 s blotting time, 0 blotting force). Data collection was performed at a CRYO ARM™ 200 (Jeol) equipped with Gatan K2 summit direct detector. A total of 8212 movies in (electron) counting mode were collected at 50 000× magnification and a pixel size of 0.968 Å/pixel. The omega energy filter was operated with a slit width of 20 eV. A total exposure of 40 e<sup>-</sup>/Å<sup>2</sup> was fractionated over 40 frames (1 e<sup>-</sup>/pixel/s) at a defocus range from -0.6 µm to -2 µm using SerialEM software (59).

#### Transcription elongation complex

100 µg of RNAP were used to assemble a TEC on a synthetic elongation scaffold as described in the section 'Elongation complex assembly for cryo-electron microscopy analysis'. TECs were applied to a Superose 6 Increase 3.2/300 column and run in cryo-buffer (100 mM HEPES/KOH pH 8.0, 100 mM NaCl, 2.5 mM MgCl<sub>2</sub>). The fractions corresponding to the polymerase peak were collected and used for vitrification (Supplementary Figure S2c, d). The sample was applied onto freshly glow discharged with EasiGlow, (TedPella) (2 × 100 s;

0.4 mbar; 15 mA) UltraAufoil R1.2/1.3 gold grids, 300 mesh (Quantifoil) and plunge-frozen into liquid ethane (Vitrobot Mark IV; Thermo Fischer Scientific; 100% humidity; 4°C; 3 s blotting time, 15 blotting force). Data was acquired on a Titan Krios G3 electron microscope equipped with a Falcon III detector using EPU software (Thermo Fisher Scientific). Movies were recorded at 75 000 $\times$  magnification (pixel size 1.0635 Å/pixel) in linear mode at a defocus ranging from  $-1.2$  to  $-2.0$   $\mu\text{m}$ . A total exposure of 86  $\text{e}^-/\text{\AA}^2$  was fractionated over 40 frames, the electron dose was 19  $\text{e}^-/\text{pixel/s}$ .

### Transcription Elongation Complex with Spt4/5

100  $\mu\text{g}$  of RNAP were used to assemble a TEC using the synthetic elongation scaffold and this sample was run on a Superose 6 Increase 3.2/300 column in cryo-buffer (100 mM HEPES/KOH pH 8.0, 100 mM NaCl, 2.5 mM  $\text{MgCl}_2$ ). 50  $\mu\text{g}$  from the most concentrated fractions were collected and used to crosslink the TEC to Spt4/5 using a BS<sup>3</sup> crosslinker (description of the crosslinking procedure is outlined in the section ‘Elongation complex assembly for cryo-electron microscopy analysis’). BS<sup>3</sup>-crosslinking can stabilize the association for transcription factors, without influencing the architecture of their complexes (60). Elution fractions from a second size exclusion chromatography run corresponding to the RNAP peak were collected and used for vitrification (Supplementary Figure S3). The sample was applied onto freshly glow discharged with EasiGlow, (TedPella) ( $2 \times 100$  s, 0.4 mbar; 15 mA) UltraAufoil R1.2/1.3, 300 mesh gold grids (Quantifoil) and plunge-frozen into liquid ethane (Vitrobot Mark IV; Thermo Fischer Scientific; 100% humidity; 4°C; 3 s blotting time, 15 blotting force). Data collection was performed at a CRYO ARM™ 200 (Jeol) equipped with Gatan K2 summit in (electron) counting mode using SerialEM software (59). Movies of 40 frames were acquired at 50 000 $\times$  magnification (pixel size 0.968 Å). The electron dose was 1  $\text{e}^-/\text{px/s}$  with a total of  $\sim 40$   $\text{e}^-/\text{\AA}^2$ . The defocus span was of  $-0.6$   $\mu\text{m}$  till  $-2$   $\mu\text{m}$  with 0.2  $\mu\text{m}$  steps.

## Data processing and model building

### Apo RNA polymerase

8212 movies of 40 frames each were collected and processed in RELION 4.0 (for the processing tree see Supplementary Figures S4, S5) (61). The movie frames were aligned, and dose weighted using RELION’s implementation MotionCor2 (62) and Contrast Transfer Function (CTF) estimation using GCTF. 6482 micrographs were selected for further evaluation based on astigmatism values, ice contamination or defocus values. Particles were first picked using Laplacian-of-Gaussian algorithm and an initial 2D classification was run. The best classes were used for Topaz (63) training and picking. 2D classification was performed twice and the best classes were selected for unbiased 3D reconstruction using 3D Initial Model function in RELION with four classes. After sequential 3D classification and filtering unaligned particles or contaminants, two classes corresponding to two different conformations: expanded conformation and super-contracted conformation with Spt4/5, were processed further separately. 196 232 particles for the expanded conformation were CTF refined and polished with Bayesian polishing function leading to a 3.5 Å map which was further classified without im-

age alignment, using a mask focused onto the clamp and stalk domain. 161 531 particles corresponding to the best aligned classes were further refined yielding a 3.2 Å resolution map.

Concerning the contracted conformation with Spt4/5, a mask incorporating the clamp extra-density was created and 3D classification without image alignment was performed. Particles corresponding to the best classes were selected, CTF corrected and polished. A final map of 3.4 Å was obtained. Final resolution was calculated using a cut-off at 0.143 gold standard in 3DFSC (64) online tool. To better describe the conformational heterogeneity of the polymerase, we re-analyzed the dataset using cryoSPARC (65). Contamination and badly aligned particles were removed in two rounds of 2D classification. Particles were refined on a newly generated initial model and 3D classification separated polymerases that were occupied by Spt4/5 bound from unbound polymerases. Both sets of particles were subsequently refined and the conformational variability was evaluated by 3D variability analysis (3DVA). For the unbound RNAP, we firstly discarded putative dimeric RNAPs, using 2D and 3D classification, to avoid bias of multimeric enzyme conformation. Principal component analysis indicated flexibility of the clamp domain and polymerase contraction (see Supplementary Figure S5 and Supplementary Movie S1). Particles were sorted into clusters along one component, a subset of 111 194 particles was selected and refined, resulting in a 3.0 Å reconstruction of the contracted apo enzyme (in contracted conformation, Figure 1B).

3DVA of the Spt4/5 bound polymerase indicated flexibility in the jaw and lobe domain, inducing polymerase contraction (Supplementary Movie S2). The lobe domain appears to contact Spt5 and might stabilize its association. We could not classify a distinct state, EM density in this part was still poor and we refrained from modelling this interaction.

### Transcription elongation complex

The data collected was 2952 movies with 40 frames each, which were processed in RELION 3.1 and RELION 4.0 beta (processing tree Supplementary Figure S6). The frames were aligned, and dose weighted as described before. 1229339 particles picked with crYOLO (66) were classified in 2D. Around 90% of the starting particles were used for generating 3D model using the initial model option in RELION. Sequential 3D classifications were performed to separate best aligned particles from contaminations. Finally, 476 655 particles were CTF refined and polished to a final resolution of 3.0 Å determined with 3DFSC online tool at 0.143 cut-off gold standard.

### Transcription elongation complex bound to Spt4/5

6076 movies of 40 were corrected as described before. Particles were first picked with Laplacian-of-Gaussian algorithm to generate 2D classes to be used for a second round of particle picking using a reference-based picking algorithm (for the processing tree see Supplementary Figure S7). Picked particles were extracted with a binning factor of 4 $\times$  and 2D classification was done. A first 3D model was generated with the initial model function in RELION, followed by sequential 3D classification and finally 322 138 particles were converted to the original pixel size. CTF correction and Bayesian polishing was performed followed by another round of CTF refinement. A 3.3 Å resolution map was generated which was further

classified without image alignment with a mask focused to the clamp and stalk domain eliminating particles not presenting Spt4/5 density on the clamp domain. The selected classes were classified again without image alignment enlarging slightly the mask. Particles corresponding to the best class were refined yielding a 3.45 Å resolution map determined as previously. To improve local sharpening for easier model interpretation, LocScale (67) software from CCP-EM (68) was used on this map.

Model building was performed on the TEC map with a homology model created on SwissModel (69) (Expasy) using *Thermococcus kodakarensis* RNAP (PDB: 6kf9) as template model. The homology model was then fitted into the map using WinCoot 0.9.8.1 (70) and Isolde (71) from ChimeraX 1.5 (72). The models were refined in PHENIX 1.20.1-4487 (73).

### Sample preparation for mass spectrometry analysis

*P. furiosus* pMUR1 and HisOnly strains were resuspended from a fresh capillary from the Regensburg Bacterial Bank and cultivated in 40 ml  $\frac{1}{2}$  SME medium supplemented with 0.1% (w/v) yeast extract, 0.1% (w/v) peptone, 0.4% (w/v) sodium pyruvate and 0.05% (w/v) Na<sub>2</sub>S and using a N<sub>2</sub> gas phase (80 kPa pressure). The medium was furthermore supplemented with 10 µM simvastatin to restrict the growth to the *P. furiosus* mutant that carries the genetically modified *rpo3* locus. The ‘His-tag only’ mutant was a genetically engineered strain with the strategy described in Waegel *et al.* (53) and was used as a negative control since it contains a plasmid that expresses only the 6xHis-tag. The re-cultivated cultures were grown at 85°C for approximately 72 h. Five cultures of each mutant were inoculated in the same medium in the absence of simvastatin and grown at 95°C for 12 h. The cultures were pelleted at 4000 g for 30 min at 4°C. The cells were plunge-frozen in liquid nitrogen and kept at -80°C until further use. The cells were resuspended in 700 µl homogenizing buffer (100 mM HEPES/KOH pH 8.0, 600 mM NaCl, 20% (v/v) glycerol, 20 mM imidazole, 2.5 mM MgCl<sub>2</sub>). The cells were lysed by sonication using Bandelin Sonopuls HD 2070 at 25% amplitude for three min. The sonication was repeated six times. Cell debris was removed by ultracentrifugation at 20 000 g for 1 h at 4°C using a Beckman Optima MAX Ultracentrifuge and a TLA-55 rotor. The supernatant was carefully collected and loaded onto 30 µl MagneHis™ beads (Protein Purification System, Promega) and incubated for 1 h, gently rotating. Subsequently, the supernatant was removed from the beads. Four washing steps were performed (three min each) using the homogenizing buffer. Finally, the beads were incubated for 15 min with 50 µl elution buffer (100 mM HEPES/KOH pH 8.0, 600 mM NaCl, 20% v/v glycerol, 500 mM imidazole, 2.5 mM MgCl<sub>2</sub>) to elute the polymerase and all interacting protein partners. One preparation was used as quality control as shown in the 4–20% gradient SDS-PAGE in [Supplementary Figure S8](#). The four replicates were precipitated in 250 µl 99.5% acetone (acetone was pre-cooled to -20°C) by incubating the samples at -20°C for 1 h. The precipitated protein mixture was centrifuged at 18 000 g for 1 h at 4°C. The supernatant was carefully removed, and the pellet was resuspended in 5 µl of 6× protein loading buffer solution (350 mM Tris/HCl pH 6.8, 30% glycerol (v/v), 10% sodium dodecyl sulphate (w/v), 0.6 M DTT, 0.012% bromophenol blue (w/v)).

## Mass spectrometry analyses

### Sample preparation

Samples were reconstituted in 1 × NuPAGE LDS Sample Buffer (Invitrogen), separated on 4–12% NuPAGE Novex Bis-Tris Minigels (Invitrogen) and stained with Coomassie Blue. Lanes were cut into 23 equidistant slices irrespective of staining. After washing, gel slices were reduced with dithiothreitol (DTT), alkylated with 2-iodoacetamide and digested with Endopeptidase Trypsin (sequencing grade, Promega) overnight. The resulting peptide mixtures were then extracted, dried in a SpeedVac, reconstituted in 2% acetonitrile/0.1% formic acid (v:v) and prepared for nanoLC-MS/MS as described previously (74).

nanoLC-MS/MS analysis for mass spectrometric analysis samples were enriched on a self-packed reversed phase-C18 precolumn (0.15 mm ID × 20 mm, Reprosil-Pur120 C18-AQ 5 µm, Dr Maisch, Ammerbuch-Entringen, Germany) and separated on an analytical reversed phase-C18 column (0.075 mm ID × 200 mm, Reprosil-Pur 120 C18-AQ, 1.9 µm, Dr Maisch) using a 46 min linear gradient of 4–35% acetonitrile/0.1% formic acid (v:v) at 300 nl min<sup>-1</sup>. The eluent was analyzed on Q ExactiveHF hybrid quadrupole/orbitrap mass spectrometer (Thermo Fisher Scientific, Dreieich, Germany) equipped with a FlexIon nanospray source and operated under Xcalibur 2.5 software using Top30 data-dependent acquisition methods, respectively. One full MS scan across the 350–1600 *m/z* range was acquired at a resolution setting of 120 000, with an AGC target of 1\*10e6 and a maximum fill time of 60 ms. Precursors of charge states 2–6 above a 4\*20e4 intensity threshold were then sequentially isolated at 1.6 FWHM isolation window, fragmented with nitrogen at a normalized collision energy setting of 30%, and the resulting product ion spectra recorded at a resolution setting of 15000 FWHM, AGC targets of 1\*10e5 and a maximum fill time of 60 ms. Selected precursor *m/z* values were then excluded for the following 25 s.

### Data processing

Raw data were processed using Mascot 2.3.02 (Matrix-science, London, UK) using standard search parameters (Trypsin/P, maximum two missed cleavages, fixed modification C-carbamidomethyl, variable modification protein N-terminal acetylation and M-oxidation, 20 ppm precursor and fragment tolerance, 1% FDR as established by a target/decoy search against a forward/reverse concatenated database). Data were searched against the UniProtKB *Pyrococcus furiosus* database (7054 entries). Search results were then combined and analyzed for normalized spectral counting quantification in Scaffold 5.2.1 software (Proteome Software, London, UK).

### Mass spectrometry data analysis

Preprocessing was performed to address missing values and normalize spectral counts. Briefly, zeros in the dataset were replaced with NA to accurately represent undetected proteins and log<sub>2</sub> transformation with subsequent median normalization was applied to mitigate the mean-variance relationship. Principal component analysis (PCA) was employed to identify similarities between samples and to detect outliers, leading to the removal of outliers from the analysis (specifically, replicate 4). The proDA package (<https://github.com/const-ae/proDA>, v.1.10.0) was utilized to analyze label-free mass spectrometry data. The normalized data were fitted with a linear proba-

bilistic dropout model within the proDA framework. Experimental design, including contrasts between different groups or conditions, was explicitly specified during the analysis. To identify proteins with coefficients significantly different from zero, hypothesis testing was conducted using proDA. A one-sided Wald test was employed to detect differentially abundant proteins, with multiple hypothesis testing correction using Benjamini-Hochberg (Supplementary Tables S3 and S4).

### Immunodetection

*Pfu* RNAP, Spt4/5 wild-type, Spt4/5 mutant and Spt4 and Spt5 proteins were loaded onto a 16% SDS-polyacrylamide gel and ran at 180 V for 40 min. The proteins were transferred onto Immobilon®-PSQ PVDF Membrane (Merk) using transfer buffer (25 mM TRIS, pH 8.3, 192 mM glycine, 10% methanol (v/v)). The transfer was carried for 1 h at room temperature at 90V/350 mA with constant voltage. After the transfer, the western blot sandwich was disassembled and the membrane was blocked in blocking solution (TRIS 10 mM, pH 8, NaCl 150 mM, 0.1% Tween 20 (v/v), 5% milk powder (w/v)) and incubated with rabbit anti-Spt5 primary antibody (1:1000 volume ratio) overnight at 4°C. The day after, the membrane was washed in TBS-T buffer (TRIS 10 mM, pH 8, NaCl 150 mM, 0.1% Tween 20 (v/v)) and incubated in blocking solution and chicken anti-rabbit IgG (H + L) secondary antibody, Alexa Fluor™ 647 (Invitrogen) in a 1:10 000 volume ratio. The Alexa Fluor 647 fluorescence signal was detected at a ChemiDoc™MP Gel Imaging System (BIORAD).

### In vitro transcription

The TEC was assembled using Cy5-labelled RNA as part of the elongation scaffold. The fluorophore was attached to the 5'-end of the 22 nt long RNA. RNA (100 μM) and the TS (400 μM) were annealed at a 3:1 molar ratio, incubated at 93°C for 3 min and slowly cooled down in a C1000 Touch™ Thermal Cycler (Bio-Rad). The pre-annealed RNA-TS (32 μM) was mixed with 50 μg of *Pfu* RNAP (4.3 μM) in a 3:1 molar ratio respectively, and the mixture was incubated at 80°C for 15 min at 80°C. Finally, the NTS was added to the mixture with a 3:1 molar ratio with respect to the RNAP concentration and incubated again at 80°C for 10 min. To separate the TEC from the excess of nucleic acids, we used a NICK column (GE Healthcare), and fractions were collected manually. The presence of the complex was assessed by 4–15% gradient SDS-PAGE using Mini-PROTEAN® TGX™ Precast Protein gels (BIORAD). The concentration of the RNAP in these fractions was determined by absorption spectroscopy at a wavelength of 280 nm using NanoDrop™ One/OneC Microvolume UV-Vis spectrophotometer (Thermo Fisher Scientific) and the most concentrated fraction (0.6–0.7 mg/ml) was used for transcription assays. The transcription reactions were performed as a time course in the absence of Spt4/5, with Spt4/5 wild-type and the Spt4/5 mutant. The first reaction mixture contained 1 μl of TEC, 3 μl 5× EMSA buffer (200 mM HEPES/KOH pH 7.4, 300 mM NaCl, 2.5 mM MgCl<sub>2</sub>·6 H<sub>2</sub>O, 0.5 mM EDTA, 0.1 mg/ml BSA), and DEPC water to arrive to a final volume of 13 μl. 1.5 μl from the nucleotide mix (ATP/CTP 1 mM, UTP 0.02 mM stock solution) was added just before incubation at 80°C. The reaction was stopped through ion chelation by the addition of 0.5

μl EDTA (0.5 M stock solution) after 30 s, 1 min, 5 min, 10 min and 30 min. The samples were denatured by the addition of 15 μl of 100% formamide buffer containing 0.05% bromophenol blue and incubation at 95°C for 10 min. The other two reactions were prepared the same way with the difference that 1 μl of Spt4/5 wild-type or Spt4/5 mutant was added, corresponding to a final concentration of 1.7 μM. For each reaction, a control was also prepared which consisted of the same reaction mixture, but the nucleotides were omitted from these samples. The elongation product was evaluated by running an 8 M urea-PAGE (with 20% v/v acrylamide concentration) at 50°C constant temperature and 25 W. The gel plates were 16 × 20 cm and 0.4 mm width, and the gel was loaded with 8 μl of the denatured sample. Detection of the Cy5 fluorophore was carried out using the ChemiDoc™MP Gel Imaging System (BIORAD), and band quantification was also performed using the software ImageLab (BIORAD). To evaluate the Spt4/5 effect on elongation the run-off product from the reaction without Spt4/5 from the 30 min sample was equalled to 100% band intensity in each gel. The experiment was performed 4 times.

### Single-molecule FRET studies

The production of double-labelled RNAP from *Methanocaldococcus jannaschii* (*Mja*) carrying a fluorescent donor and acceptor dye pair as well as a biotin label in subunit Rpo3 was described in detail elsewhere. The expression and purification protocol to produce recombinant *Mja* Spt4/5 was reported in detail in the following publication (20). Briefly, measurements on immobilized RNAP molecules were carried out using a homebuilt PRISM-TIRF setup based on an Olympus IX71 microscope equipped with alternating laser excitation. All measurements were carried out at room temperature (21°C) using a measurement chamber passivated with PEG. Passivated quartz slides and measurement chambers were prepared according to Schulz *et al.* (19) and Roy *et al.* (75). For excitation, a yellow laser (568 nm; Coherent Sapphire 100 mW) was used for excitation of the donor, and a red laser (639 nm; Topica iBeam Smart 150 mW) was used for direct excitation of the acceptor, respectively. The fluorescence was collected using a 60x Olympus 1.20 N.A. water-immersion objective. Emission light was split by wavelength using a dichroic mirror (640 DCXR, Chroma Technology) into two detection channels that were further filtered with two bandpass filters (Semrock BrightLine 582/75 and Semrock Brightline 609/54) in the orange channel and one longpass filter (647 nm Semrock RazorEdge) in the near-infrared detection range. Both detection channels were recorded by one EMCCD camera (AndorIX on X3, preGain 5.1, gain 250, frame rate 10 Hz) in a dual view configuration (TripleSplit, Cairn UK) and the acquired data were analyzed by custom-made software based on LabVIEW 2012 64bit (National Instruments). Data analysis was furthermore carried out using the software Origin. The resulting histograms were fitted either with a single or double Gaussian fit, and the mean FRET efficiency and the standard error was determined from the fit. Spt4/5-RNAP complexes were pre-formed at 65°C for 20 min prior to immobilization on the streptavidin-coated surface and subsequent microscopy. Similarly, RNAP apo enzyme was incubated for 20 min at 65°C without the addition of Spt4/5 prior to single-molecule measurements.

## Results

### Cryo-EM reconstructions of apo RNA polymerase from *P. furiosus* show the clamp/stalk module in expanded and contracted states

In this work, we purified the native archaeal RNAP from the hyperthermophilic euryarchaeon (Methanobacteriota) *Pyrococcus furiosus* (Supplementary Figure S1). To this end, a mutant *P. furiosus* strain constitutively expressing C-terminally Strep- and His-tagged RNAP subunit Rpo3 from its native genomic locus was created, allowing affinity purification of the cellular archaeal RNA polymerase (Supplementary Figure S1). *Pfu* RNAP is used as a standard model system for archaeal transcription and shows high sequence identity (83.8%) with the archaeal RNAP from *Thermococcus kodakarensis*, of which structures were analyzed before (46,49). Cryo-EM data was collected on a CryoARM 200 (JEOL) equipped with a Gatan K2 direct electron detector and an omega filter and processed using the RELION (61) suite. Single-particle reconstructions of the apo form of the *Pfu* RNAP were determined at final resolutions of up to 3.2 Å (Figure 1, Supplementary Figures S4, S5, Supplementary Table S2). Damaged, oligomerized and undefined particles were removed during data processing (Supplementary Figure S1). *Pfu* RNAP is composed of 11 subunits devoted to catalysis (Rpo1N, Rpo1C and Rpo2), assembly of the RNAP (Rpo3, Rpo10, Rpo11 and Rpo12), and auxiliary functions (Rpo4, Rpo7, Rpo5 and Rpo6) (Supplementary Figure S9). The reconstructions reveal well-defined cryo-EM density for all 11 *Pfu* RNAP subunits, allowing model building (Supplementary Figure S10) starting from *Thermococcus kodakarensis* RNAP models (46,49). A reconstruction containing 161 531 particles shows the RNAP in an open clamp conformation characterized by an expanded central DNA-binding cleft spanning ~54.3 Å from the clamp core helices to the protrusion domain (Supplementary Figure S11a). The DNA-binding cleft is not occupied by any density, indicating that a complete removal of residual nucleic acids during purification was successful, and an inactive apo-RNAP reconstruction was obtained. Central functional elements such as the bridge helix and the catalytic Mg<sup>2+</sup> ion are well ordered and show defined density, while residue positions in the trigger loop remain flexible (Figure 1A, D). However, local resolution estimation in RELION indicated dynamic conformations of the stalk domain (subunits Rpo4/7) as well as the clamp head, clamp core, and the jaw domain within subunit Rpo1C. To analyze the structural heterogeneity of these domains, we employed the variability analysis included in CryoSPARC (65), confirming the dynamics of the clamp/stalk (Supplementary Movie S3) and jaw regions (Supplementary Movie S1). Furthermore, this allowed the reconstruction of an additional contracted RNAP conformation at a resolution of 3.0 Å (Figure 1B and Supplementary Figure S5). These results indicate that the *Pfu* RNAP exists in a continuum of structural conformations within an expanded and a contracted boundary in its apo form *in vitro*.

This is well in line with the results of single-molecule FRET studies using the recombinant archaeal RNAP from the related archaeal organism *Methanocaldococcus jannaschii*, showing that in the apo state the *Mja* RNAP can adopt an expanded and contracted conformation, with additional molecules adopting a conformation in between, suggesting

that also the *Mja* RNAP is in a dynamic equilibrium between both states (Figure 2D).

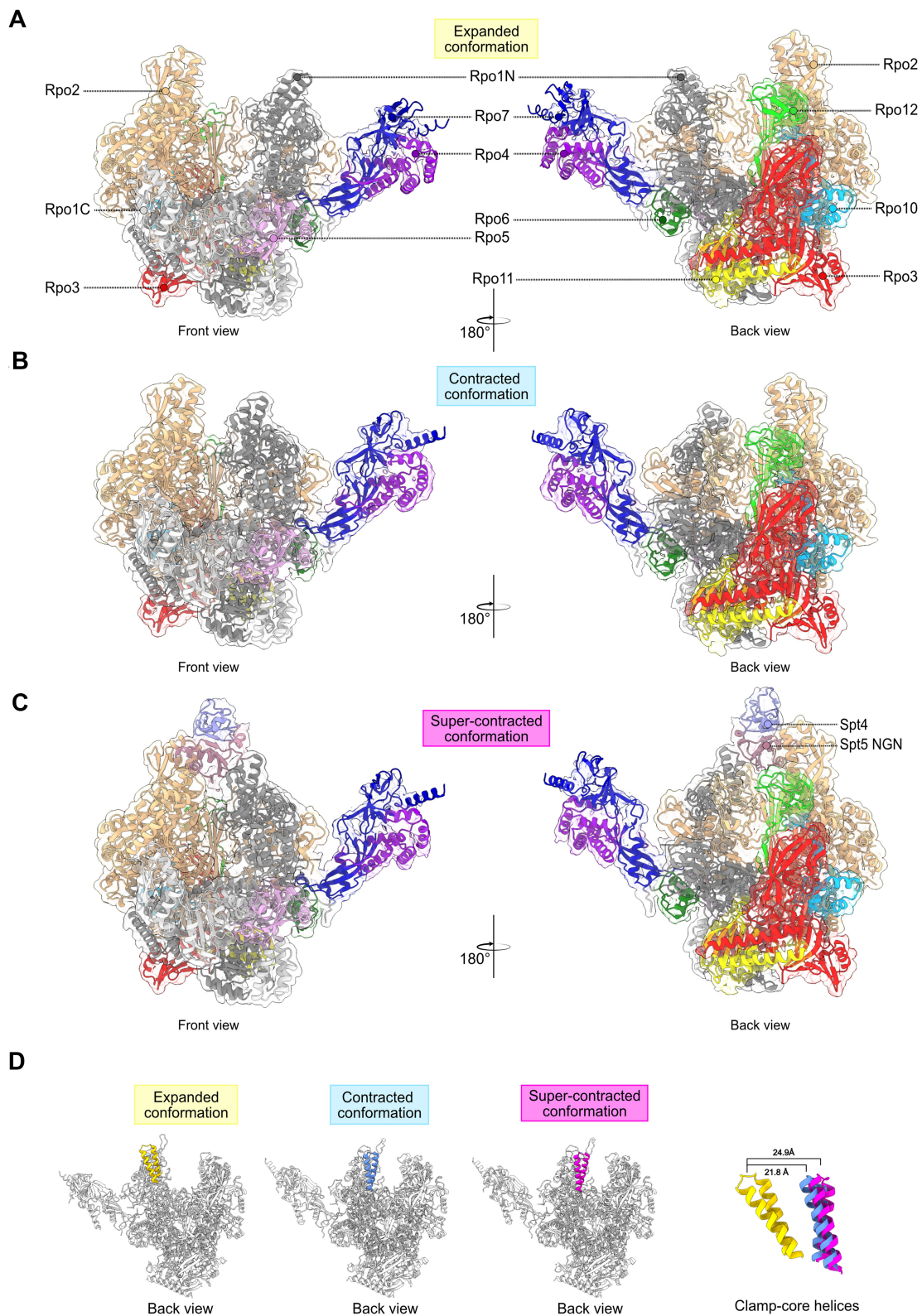
### Spt4/5 binds the RNA polymerase in the absence of nucleic acids and induces a conformational change

In the same cryo-EM dataset of the apo *Pfu* RNAP, about half of well-defined particles showed additional cryo-EM density spanning the DNA-binding cleft between the clamp core helices and the protrusion domain. Following 3D classification and refinement in RELION, a 3D reconstruction of these particles at an overall resolution of 3.4 Å was obtained (Figure 1C, Supplementary Figure S4 and Supplementary Table S2). Using the model of a recombinant clamp-Spt4/5 crystal structure as reference (21), we were able to unambiguously attribute the additional density to Spt4/5. The position overlaps with interpretations of low-resolution EM envelopes. Density for the Spt5 KOW domain was not observed in our reconstructions, indicating that the domain is flexibly linked to the NGN domain in this state and the KOW domain has rotational freedom. Local resolution differences (Supplementary Figure S5) and flexibility analysis in CryoSPARC indicate that the position of Spt4/5 is not rigidly defined (Supplementary Movie S2). The RNAP appears to be in a ‘super-contracted’ conformation when bound by Spt4/5, meaning that the clamp domain closes beyond the most contracted apo *Pfu* RNAP conformation we observed (Supplementary Movies S1, S4). In the super-contracted conformation, the stalk sub-complex and the jaw domain contract alongside the clamp domain. Furthermore, the Spt5-NGN domain reaches so far across the DNA-binding cleft that it contacts the lobe in addition to the protrusion domain. Note that the DNA binding channel of the apo *Pfu* RNAP is not accessible for DNA and that initiation factor binding sites are occluded in the apo *Pfu* RNAP-Spt4/5 configuration, preventing productive initiation or elongation (Supplementary Movie S2). Consequently, Spt4/5 would have to dissociate to allow the loading of the DNA during the initiation phase of transcription.

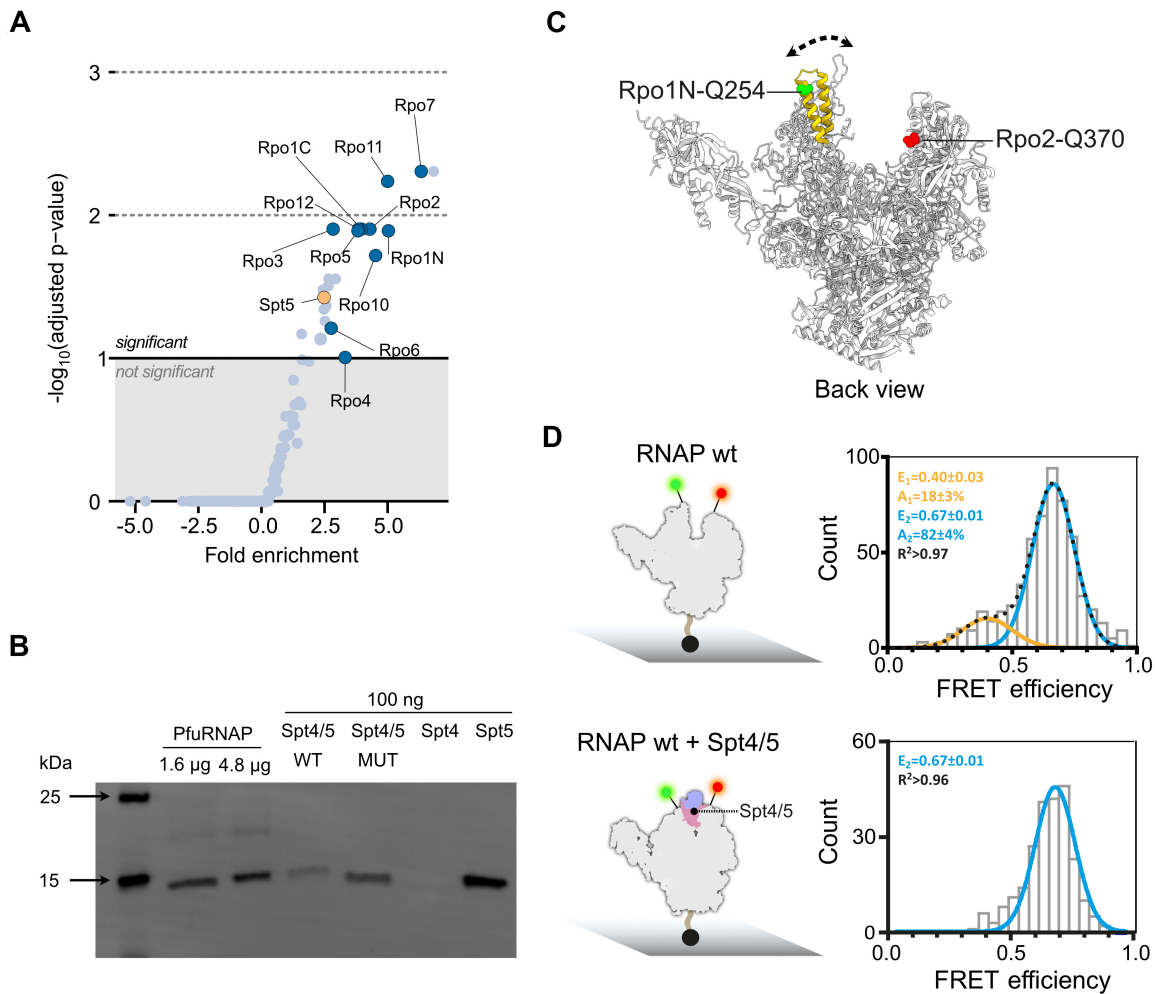
To verify that endogenous Spt4/5 indeed co-purifies with *Pfu* RNAP from cellular extracts, we performed mass spectrometry (MS) analysis of the affinity purified *Pfu* RNAP (Supplementary Figure S8). In addition to all RNAP subunits, Spt5 was also enriched in these samples (Figure 2A). Spt4 was slightly less enriched, likely owing to the small size of the protein. In addition to MS analysis, the co-purification of Spt4/5 with *Pfu* RNAP was confirmed by Western Blot analysis using an antibody directed against *Pfu* Spt4/5 (Figure 2B).

We next asked whether the different conformational states of the *Pfu* RNAP and *Pfu* RNAP in complex with Spt4/5 can also be monitored in solution. Complementing the cryo-EM analysis, we made use of our previously established single-molecule FRET assay using the RNAP of the closely related archaeon *M. jannaschii*. This assay reports on the distance and the distance changes between a fluorescent donor-acceptor dye pair engineered into two residues that are located opposite of each other across the DNA-binding cleft (Figure 2c, residue 257 in Rpo1N and residue 373 in Rpo2C, respectively numbering for the RNAP of *M. jannaschii*). The *Mja* RNAP is found in an expanded (low FRET with mean FRET efficiency of  $E = 0.40$ ) and contracted state (high FRET with mean FRET efficiency of  $E = 0.67$ ) in solution (Figure 2D) reflecting the conformational states found in our cryo-EM





**Figure 1.** Structure of the euryarchaeal RNA polymerase apo enzyme and RNA polymerase-Spt4/5 complex from *Pyrococcus furiosus*. **(A)** Structural model fitted into the corresponding refined density map of the apo *Pfu* RNAP with an open clamp in two different views with the corresponding subunit denomination. **(B)** Structural model of apo *Pfu* RNAP and corresponding refined density map in a closed clamp conformation in two different views. **(C)** Structural model of the *Pfu* RNAP with Spt4/5 fitted into the refined density map shows that the clamp domain adopts a closed conformation. **(D)** Direct comparison of *Pfu* RNAP structures with open and closed clamp and close-up view of the clamp coiled coil helices that undergo a movement of 21.4 Å upon transition from the open to closed state. Structures were globally aligned in ChimeraX, contraction distance was measured between the C $\alpha$  positions of Proline 261 within subunit Rpo1N in the three models.

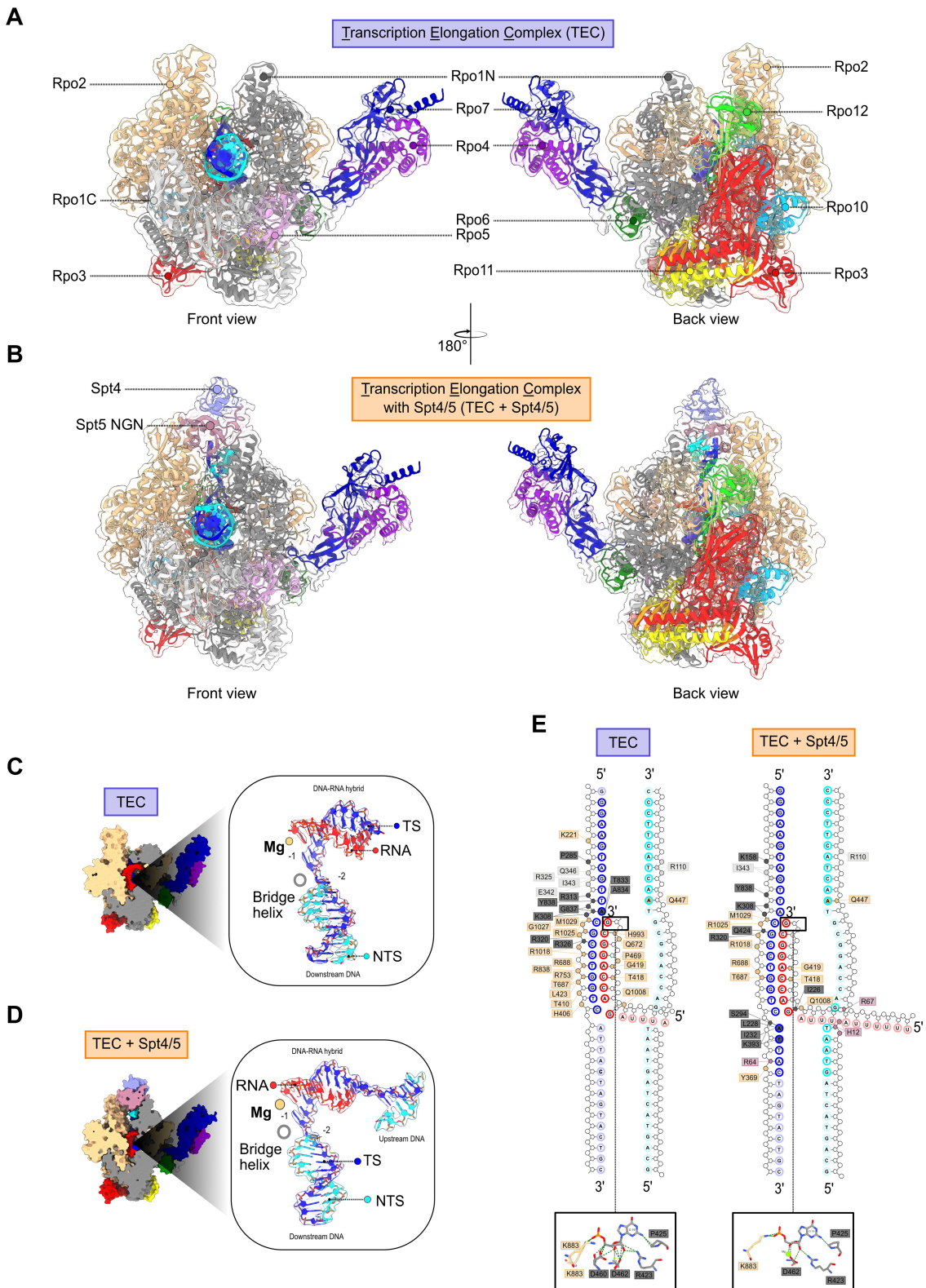


**Figure 2.** Accessing the *Pfu* RNA polymerase-Spt4/5 interaction using mass spectrometry and single-molecule FRET measurements. **(A)** Volcano plot of mass spectrometry analysis conducted for the *P. furiosus* strain that encodes a Strep-HIS-tagged variant of the subunit Rpo3 with respect to the negative control ('His-tag only' mutant). Fold enrichment over background was calculated using a one-sided Wald test with multiple testing correction using Benjamin-Hochberg. RNAP subunits and Spt5 are color-coded in blue and yellow, respectively. **(B)** Western blot analysis of samples derived from affinity-purified *Pfu* RNAP from a cell lysate from a *P. furiosus* variant that expresses Strep-HIS-tagged subunit Rpo3. Recombinantly expressed and purified wild-type or mutant Spt4/5<sup>H12A/R64A/H65A/R67A</sup> (100 ng) as well as the individual recombinant proteins Spt4 and Spt5 (100 ng) were loaded as reference. For detection, an antibody raised against the elongation factor Spt4/5 was used. **(C)** Structure of the 12-subunit archaeal RNA polymerase from *Pyrococcus furiosus* (PDB: 8ORQ) with an open clamp (yellow). The archaeal RNA polymerase from *Methanocaldococcus jannaschii* was site-specific labelled with a donor (DyLight550) and acceptor fluorophore DyLight650 in subunit Rpo1N\*<sup>E257</sup> (green) and the protrusion domain (subunit Rpo2C\*<sup>Q373</sup>, red), respectively. For immobilization purposes a biotin was coupled to a native cysteine in subunit Rpo11 (black sphere protruding from the RNAP core). **(D)** FRET efficiency histograms show smFRET efficiencies determined for the wildtype *Mja* RNA polymerase (wt *Mja* RNAP) and in complex with transcription factor Spt4/5 (wt *Mja* RNAP + Spt4/5). All measurements were performed in the absence of nucleic acids. The histograms were fitted with a single or double Gaussian function and the mean FRET efficiencies ( $E$ ), the percentage distribution of the populations ( $A$ ) and the coefficient of determination ( $R^2$ ) are given with standard errors in the histograms.

studies. Our data furthermore reveal that the addition of Spt4/5 shifts the equilibrium towards the population with high FRET efficiency (Figure 2D). Even though we observed a further closure of the clamp for the apo *Mja* RNAP in the presence of Spt4/5 yielding the super-contracted *Pfu* RNAP state, we do not observe a shift in mean FRET efficiency, which is the same as for the unliganded apo *Mja* RNAP ( $E = 0.67$ ). However, a distance change of 2 Å between the contracted and super-contracted state would only result in a marginal change in FRET efficiency of 0.03, which is too small to be reliably detected. Taken together, these data further support the finding that binding of Spt4/5 to the apo RNAP is possible even in the absence of DNA and that Spt4/5 induces conformational changes in the apo RNAP.

### The structure of an archaeal transcription elongation complex is similar to eukaryotic counterparts, with some exceptions

To investigate the three-dimensional architecture of the archaeal TEC, *Pfu* RNAP was recruited to a synthetic elongation scaffold following protocols described for TECs formed with the *Mja* RNAP (46,49) and subjected to single-particle cryo-EM analysis (Figure 3A, Supplementary Figure S2c,d, S6, Supplementary Table S2). Following processing and 3D reconstruction in RELION (61), we obtained a cryo-EM map of the *Pfu* TEC at an overall resolution of 3.0 Å. In addition to the *Pfu* RNAP subunits, the density allowed us to build a molecular model for the downstream DNA double helix and the DNA-RNA hybrid, enabling investigations of nucleic



**Figure 3.** Structure of the archaeal transcription elongation complex in presence and absence of Spt4/5. Structural models fitted into the refined density map of the *Pfu* TEC in (A) the absence and (B) presence of Spt4/5 in two views. (C) Schematic representation of the nucleic acid scaffolds used for the TECs formed in the absence (top) and presence (bottom) of Spt4/5. Template strand colored in blue, non-template strand in cyan, RNA in red. Interactions with amino acids side chains with the nucleic acid scaffold are indicated. Coloring of the RNA polymerase subunits as in (A). (D) Nucleic acid strand model in the TEC and the TEC in the presence of Spt4/5. The magnesium cation in the metal A position is shown as orange ball. The position of the bridge helix is indicated with a grey sphere. Part of the downstream DNA can be traced when Spt4/5 is part of the TEC. (E) Schematic representation of *Pfu* RNAP and nucleic acids contacts calculated in ChimeraX in the TEC in absence (left) and in presence (right) of Spt4/5. The colours of the amino acid boxes correspond to the subunit color-coding which was used in (A) and (B). Full colour indicates high confidence assignments within the 3D maps, whereas faint-coloured areas are the parts of the scaffold that could not clearly be assigned from the reconstruction.

acid-protein interactions within the *Pfu* TEC (Figure 3C, E). The model includes nine ribonucleotides of the RNA, but local resolution at the areas of the RNA-exit channel and within the center of the transcription bubble prevented further RNA and non-template strand modelling. Comparable to other RNAP-TEC reconstructions (8,76) and crystal structures, the downstream dsDNA double helix was well-defined, whereas the upstream dsDNA region showed higher flexibility. The primary contacts between the *Pfu* RNAP and nucleic acids are not sequence-specific but involve several basic residues and the DNA phosphate backbone (Figure 3C). The *Pfu* RNAP is found in the pre-translocated state as indicated by the catalytic magnesium contacting the +1 ribonucleotide. The bridge helix separates position -1 from -2 on the DNA template strand and is found in a fully helical ‘unstretched’ state. Overall, the structure of the elongation complex shows high similarity to the three-dimensional organization of the eukaryotic TEC containing RNAP II (77) (Supplementary Figure S12). This includes the clamp/stalk module that adopts the contracted state, securing the downstream DNA in the DNA binding channel (Figure 3, Supplementary Movie S3). However, a conserved contact of the downstream DNA duplex with subunit Rpb5, as observed for eukaryotic TECs, is not reflected in the archaeal TEC because the N-terminal domain of eukaryotic Rpb5 is not conserved in Archaea.

Structural comparison of the *Pfu* TEC reconstruction with the Spt4/5-engaged apo polymerase showed many common features, such as the overall contracted RNAP state with a closed clamp domain, contracted stalk sub-complex and jaw domain (Figure 4A). This leads to the conclusion that the RNAP conformation would remain identical when Spt4/5 subsequently engages with the TEC. To test this hypothesis, we again employed single-particle cryo-EM to determine a reconstruction of the Spt4/5-decorated *Pfu* TEC.

### Archaeal Spt4/5 engages transcribing and non-transcribing RNAP in different orientations

To determine the structure of the archaeal TEC in complex with Spt4/5, the complex was formed by adding an excess of recombinant Spt4/5 to a pre-formed *Pfu* TEC followed by stabilization using chemically crosslinking with BS<sup>3</sup> prior to size exclusion chromatography to remove incompletely assembled particles (Supplementary Figures S3, S7). Here, we used an elongation scaffold with a longer RNA since structural studies of RNAP II in complex with Spt4/5 (78,79) showed contacts of Spt5 with the emerging RNA via its KOW domains. Again, we collected cryo-EM data followed by processing and single-particle analysis (Supplementary Figures S2, S7). This way, a reconstruction at an overall resolution of 3.4 Å resolution was obtained, allowing the construction of a molecular model of the Spt4/5-containing *Pfu* TEC (Figure 3B and Supplementary Figure S7, Supplementary Table S2). Even though the local resolution is reduced in the region where Spt4/5 interacts with the *Pfu* RNAP clamp, an adaptation of the previous model showing Spt4/5 in contact with the contracted apo *Pfu* RNAP (see above) allowed the construction of a confident model. Nevertheless, consistent density for the KOW domain was absent, indicating that the domain is either flexibly linked to the NGN domain or subject to positional heterogeneity in the archaeal TEC. Contacts between the Spt5 NGN domain and the *Pfu* RNAP clamp and protrusion domains are established, as shown

in Supplementary Figure S12a. In comparison to the previous *Pfu* TEC reconstruction, the Spt4/5-containing *Pfu* TEC shows clear density for the upstream dsDNA duplex, allowing to construct a model for this part and hinting at a stabilization of this DNA part by Spt4/5 within the TEC (Figure 3d). Furthermore, a comparison of the position of Spt4/5 within the *Pfu* TEC and in the apo *Pfu* RNAP reveals that helix 1 of Spt5 is found in a position shifted by ~6 Å (Figures 4B, C and 5A, Supplementary Figure S12). In the context of the *Pfu* TEC, Spt5 helix 1 is shifted away from the protrusion domain and thus partially retracts from the DNA-binding cleft of the RNAP. This repositioning is likely required to accommodate DNA within the central cleft. Upon partial retraction in the TEC, contacts of Spt5 with the RNAP lobe domain cannot be established any longer, preventing the transition into a super-contracted state as found in the nucleic acid-free apo *Pfu* RNAP complexed with Spt4/5.

Notably, the C-terminal region of Rpo2 (residues 1005–1024) that contacts the DNA–RNA hybrid appears to be fully flexible within the expanded *Pfu* RNAP form but adopts a fully folded and active conformation within the contracted form in a conformation that is identical to the *Pfu* TEC and Spt4/5-containing *Pfu* TEC. This stretch binds to the DNA–RNA hybrid, specifically the phosphate backbone of the template strand, in both *Pfu* TEC reconstructions.

### Spt5 apparently stabilizes the upstream DNA duplex

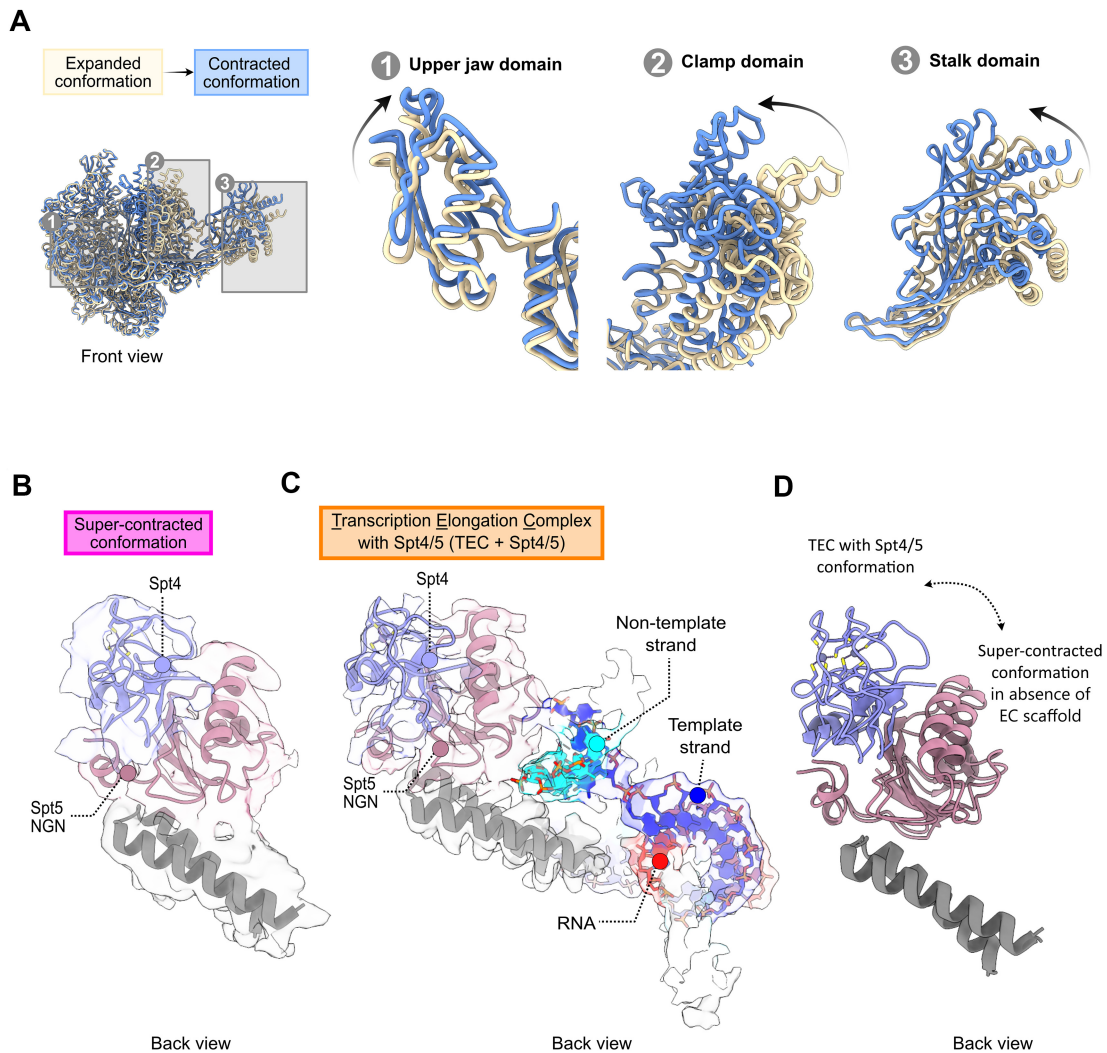
Putative contacts between the NGN domain of Spt5 and the backbone of the upstream DNA duplex in the *Pfu* TEC reconstruction were identified, involving the four residues His12, Arg64, His66 and Arg67 (Figure 5A). Although the precise sidechain orientations in this contact region remain only partially defined, our analysis indicates that the presence of Spt4/5, compared to the free *Pfu* TEC, leads to the fixation of the transcription bubble at its upstream edge, stabilizing the upstream duplex. Interestingly, the four basic Spt5 residues that contact the upstream DNA are not conserved across archaeal phyla or when compared to bacterial NusG or to eukaryotic Spt5 (Figure 5B).

To assess the significance of DNA engagement by the four identified basic residues for Spt5 function, we conducted a standardized *in vitro* transcription assay. The results demonstrate that the presence of wild-type Spt4/5 prevents pausing and stimulates the processivity of *Pfu* RNAP (Figure 5C, D, Supplementary Figure S13). However, the stimulatory effect of Spt5 was reduced in the Spt5 mutant, in which the histidine and arginine residues that contribute to the Spt5–DNA interface were mutated to alanine. It should be noted that neither heat stability nor complex formation with Spt4 was affected in this mutant (Figure 5A). Taken together, the results of this structure-function analysis indicate that the interaction between Spt5 and the upstream edge of the transcription bubble is important for Spt5 function in *P. furiosus*.

## Discussion

### Evolutionary aspects regarding clamp movement and cleft contraction

The clamp domain is a mobile element crucial to the functioning of all multi-subunit RNAPs, as shown by a wealth of structure-function and single-molecule FRET analyzes. How-

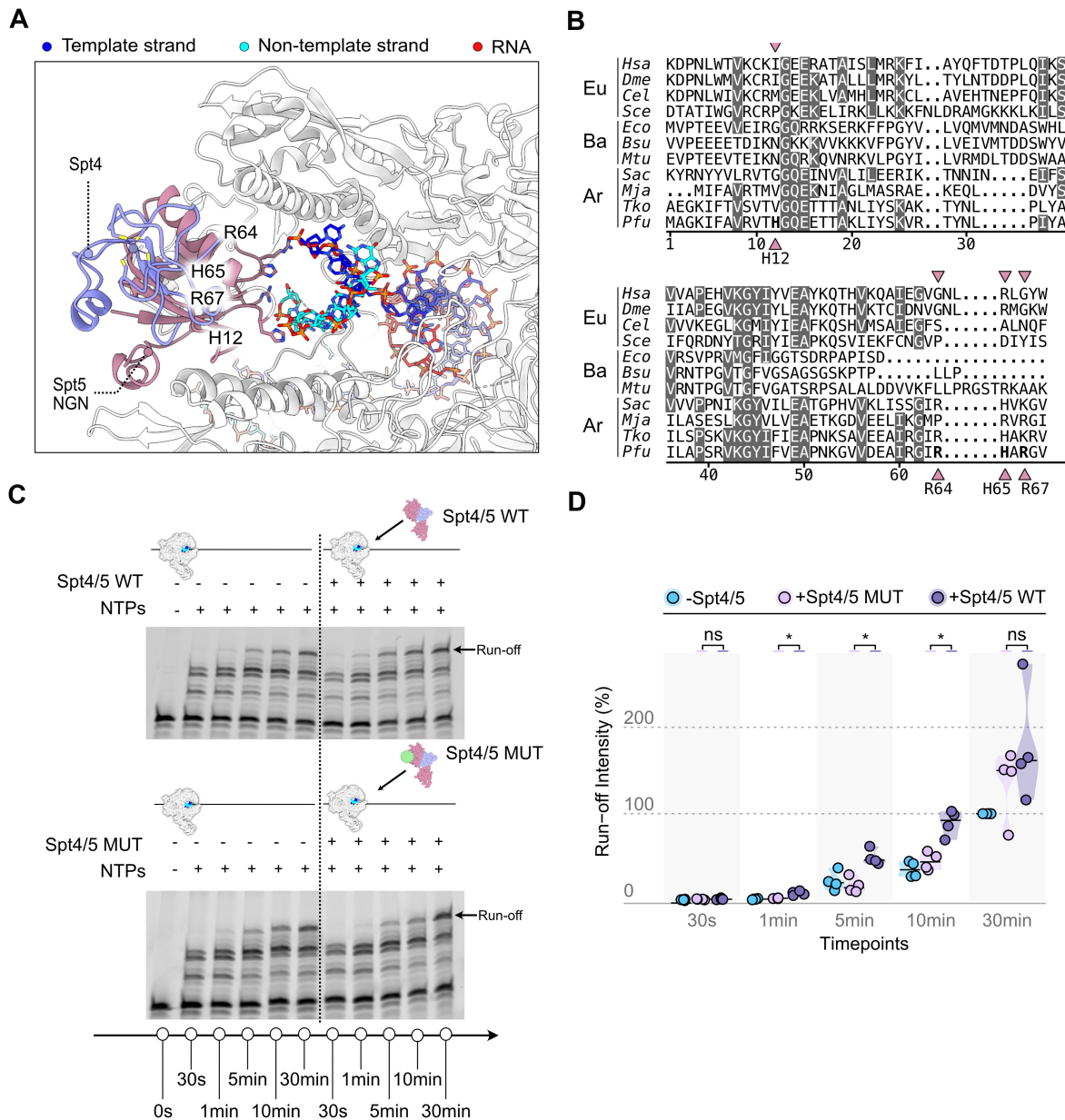


**Figure 4.** Comparison of Spt4/5 locations within the different structures. **(A)** Conformational changes between expanded and contracted conformations of the *Pfu* RNAP apo polymerase with focus on upper jaw, clamp and stalk domains. **(B)** Structural model of Spt4/Spt5 NGN domain and RNAP clamp-core helices of the super-contracted *Pfu* RNAP conformation with corresponding refined cryo-EM density. **(C)** Structural model of the Spt4/Spt5 NGN domain, clamp core helices and scaffold of the TEC with Spt4/5 with the corresponding cryo-EM density **(D)** Comparison of Spt4/5 positioning in the two conformations from (A) and (B). The models were aligned in ChimeraX via the clamp-core helices.

ever, there is no consistent picture regarding the conformation of archaeal RNAPs in their apo form, as the RNAP was found in either contracted or expanded states (36,46,49,51). In this study, we fill gaps in the understanding of the conformational status of the clamp in Archaea, bringing together previous observations of archaeal RNAP clamp flexibility in the apo state and clamp fixation upon transition to active elongation. Single-particle cryo-EM reconstructions of *Pfu* RNAP presented in this work showed the enzyme in expanded and contracted conformations that are defined by different positions of the clamp domain and the stalk sub-complex (Figure 1A, B). Three-dimensional variability analysis revealed a high degree of clamp/stalk module flexibility among all particles, which is well in line with single-molecule FRET (smFRET) measurements showing that expanded and contracted conformations can both be adopted in solution, suggesting that the clamp is fully flexible when the RNAP is not engaged with the template DNA or transcription factors. In the elongation state, the clamp closes, and the DNA-binding cleft is found in a contracted state (Figure 2A, B). In line with our findings, an-

other parallel work elucidated the structural organization of an archaeal pre-termination complex, showing the clamp in a closed conformation before termination takes place (80). It is assumed that the clamp reopens to release the DNA and RNA during termination (40,81,82).

Similarly, functional clamp movements have been observed for bacterial and eukaryotic RNAPs (6,7). Structural and smFRET analysis of bacterial RNAPs at different steps of the transcription cycle elucidated the mobile nature of the clamp domain. For example, the bacterial apo RNAP exhibits an expanded conformation with an open clamp in crystal structures and cryo-electron microscopy structures (83,84). Single-molecule FRET measurements of the apo RNAP and holoenzyme enzyme from *E. coli* moreover revealed that the clamp is conformationally flexible and populates three distinct states with the majority of molecules adopting the open clamp state (85). In bacterial transcription elongation complexes (35,86), the clamp adopts a closed state and apparently reopens during transcription termination (87). These clamp movements are of functional importance for bacterial but also



**Figure 5.** Interactions between Spt4/5 and the non-template strand within the upstream DNA duplex enhance the transcriptional activity of the archaeal RNA polymerase. **(A)** Close-up view of the interaction of Spt5 with the phosphate backbone of the non-template strand. Spt5 sidechains (H12, R64, H65 and R67) that contact the DNA are indicated. Spt5 is colored in dusky pink and Spt4 in purple. **(B)** Multiple sequence alignment of Spt5 including eukaryotic Spt5 variants (*Hsa*: *Homo sapiens*, *Dme*: *Drosophila melanogaster*, *Cel*: *Caenorhabditis elegans*, *Sce*: *Saccharomyces cerevisiae*), bacterial NusG variants (*Eco*: *Escherichia coli*, *Bsu*: *Bacillus subtilis*, *Mtu*: *Mycobacterium tuberculosis*) and archaeal Spt5 variants (*Sac*: *Sulfolobus acidocaldarius*, *Mja*: *Methanocaldococcus jannaschii*, *Tko*: *Thermococcus kodakarensis*, *Pfu*: *Pyrococcus furiosus*). Conserved residues are highlighted in grey. The residues that interact with the non-template strand in the TEC of *Pfu* are indicated with a triangle. **(C)** Transcription elongation assay using wt *Pfu* RNAP (0.6–0.7  $\mu$ g) and Spt4/5 wt or the Spt4/5 mutant (Spt5<sup>H12A/R64A/H65A/R67A</sup>) (1.7  $\mu$ M). As references, transcription assays in the absence of Spt4/5 were performed. Reactions were stopped at 0.5, 1, 5, 10 and 30 min. **(D)** Diagram showing the quantified run-off transcript levels shown in panel (C). Violin plot illustrating run-off intensity (%) over time (30 s, 1 min, 5 min, 10 min, 30 min) for a reaction without Spt4/5 (blue), for a reaction with Spt4/5 MUT (lavender), and a reaction with Spt4/5 WT (purple) samples. Individual data points are overlaid as jittered points for clarity. Medians displayed with crossbars. Statistical significance indicated by NS ( $P > 0.05$ ) and \* ( $P < 0.05$ ) based on Mann–Whitney  $U$  tests comparing + Spt4/5 MUT and + Spt4/5 WT at each time point.

archaeal transcription as an opening of the clamp promotes, for example, transcription initiation (19,88), most likely by supporting the separation of the DNA strands. Clamp flexibility was not observed for apo RNAPs from the archaeal Thermoproteota phylum. For this phylum, crystal structures (*S. shibatae* and *S. solfataricus*) (48,51) and a cryo-EM structure from *S. acidocaldarius* (36) were reported, and in all cases, a closed clamp conformation was found for the RNAP apo form. The viral RNAP-interacting protein ORF145 and TFS4 were shown to impact the *S. acidocaldarius* RNAP clamp conformation and to shift the clamp to a more open state as compared to the apo enzyme (Supplementary Figure S11b, e). Nevertheless, the crystal structure (46) of *T. kodakarensis* RNAP was resolved in an expanded conformation but was found in a contracted conformation in the cryo-EM structure (49), similar to *Pfu* RNAP apo forms. FRET and structural studies from *P. furiosus*, *T. kodakarensis* and *M. jannaschii* suggest a mechanistic conservation of a flexible clamp of the apo RNAP throughout the archaeal Methanobacteriota phylum members, which could not be verified in RNAPs from the Thermoproteota phylum to date.

Opening and closing of the clamp domain was also observed in eukaryotic RNA polymerases (40), including RNAP II of yeast and mammalian origin. However, clamp movements were documented to be more pronounced for mammalian RNAP II. For example, density for the clamp was found to be weak or absent in a cryo-EM dataset for RNAP II associated with protein 2 (RPAP2), which binds between the jaw domains of RNAP II (77,89) in absence of nucleic acids. Similarly, the clamp is highly mobile in the dimeric RNAP II structure from *S. scrofa domesticus* (90). Structures of the transcribing bovine RNAP II showed that even in the TEC, a small subfraction of the RNAPs (10%) exhibited a mobile clamp in the absence of Spt4/5. Notably, the activity of RNAP I relies on modulation of its overall contraction state, including the clamp in inactive conformations, a full opening of the clamp-stalk module occurs, even allowing RNAP dimerization (6,7). RNAP I is found in a fully contracted state upon initial transcription (91) and the clamp remains closed throughout productive elongation, as shown by single-particle cryo-EM and tomography analysis (92, 93). Similar to *Pfu* RNAP, the hybrid binding site of RNAP I is activated following contraction (94). The apo form of RNAP I may also adopt different clamp contraction states. Especially in human RNAP I, clamp flexibility is strongly pronounced, indicating a conformational continuum (95). Hence, *Pfu* RNAP and eukaryotic RNAP I appear to undergo comparable expansion/contraction movements.

### Spt4/5-engagement limits conformational flexibility of the archaeal RNAP

Interaction of bacterial NusG with DNA-free bacterial RNAP has been reported (96), but a comparable interaction between Spt4/5 and the DNA-free state of RNAP was never observed for eukaryotic RNA polymerases. This might be because a tight interaction between eukaryotic Spt4/5 and the RNAP requires the presence of the newly synthesized RNA that forms an interaction with the linker domain between the KOW<sub>4</sub> and KOW<sub>5</sub> domain of eukaryotic Spt5 (78,79). Here, we conclude that archaeal Spt4/5 binding does not strictly require KOW-interactions with the product RNA and is only possible when the RNAP adopts the closed clamp conformation (Figure 2B). For Spt5 from *M. jannaschii*, residues A4 and Y42 are of cru-

cial importance for the interaction with the RNAP (20). These residues are conserved in *P. furiosus* and similarly contribute to the interaction with the RNAP clamp via its hydrophobic pocket together with a hydrogen bond between Spt5 E78 and Rpo1N Q262, which was also previously observed in the crystal structure (21). Likely, and well in line with smFRET analyses *in vitro*, archaeal RNAP adopts various states in a dynamic equilibrium within the contracted and expanded boundaries we describe. Spt4/5 apparently selects for a specific conformational state and can only bind contracted RNAPs, resulting in super-contraction and removal from the conformational flexible pool of apo RNAPs. However, we cannot rule out that Spt4/5 binding induces the contraction of the RNAP instead.

Loading of the DNA into the binding channel requires an open state of the clamp, a structural transition supported by TFE during transcription initiation (19,49). Our cryo-EM data suggest that a subfraction of the cellular RNAP pool is associated with Spt4/5 in the absence of a nucleic acid substrate. In the *Pfu* RNAP-Spt4/5 complex, the clamp is closed, and the DNA binding channel is blocked by Spt4/5. Consequently, DNA cannot enter the RNAP in an Spt4/5 pre-engaged state. This may indicate a mechanism to prevent the binding of unspecific DNA. Similar strategies are employed in RNAP I and III using intrinsic structural elements (specifically, the C-terminus of subunit RPC7 in RNAP III (97) or the DNA-loop within the largest subunit of RNAP I (6,7)). RNAPs decorated with Spt4/5 might represent a dormant RNAP reservoir available in the cell that can quickly be activated. To activate *Pfu* RNAP molecules for promoter-directed transcription, Spt4/5 must be displaced from the RNAP. Spt4/5 is known to compete with TFE for the clamp (11,17), leading to a replacement of TFE during the transition from initiation to elongation. A reverse scenario, e.g. the displacement of Spt4/5 by TFE to prime the RNAP for transcription initiation, might allow the activation of this RNAP pool under circumstances that demand higher transcriptional output or a switch of the transcriptional program. In line with this idea, we have previously shown that TFE has a higher affinity for the RNAP in context of the PIC as Spt4/5 (17). A somewhat similar mechanism is employed by the viral protein ORF145 which occupies the DNA binding channel of the RNAP and effectively inhibits productive transcription (50). As the respective virus does not encode its own RNAP, release of ORF145 from the archaeal RNAP must take place to allow transcription of the viral genome. The mechanism by which this is achieved, however, is unknown. Compared to the RNAP-ORF145 structure, the extent of RNAP clamp closure is more pronounced in the Spt4/5-TEC (Supplementary Figure S11). For these reasons, we hypothesize that the reversible shutdown of RNAP activity by temporary fixation of the cleft contraction state may be a regulatory strategy.

### The Spt5-interface with upstream DNA in TECs is functionally important in *P. furiosus*

Bacterial NusG and archaeal Spt5 (in complex with Spt4) are well known for their function as transcription elongation factors that stimulate the processivity of the RNAP (4). The underlying mechanism for this stimulatory activity is found in the overall architecture of the TEC with Spt4/5 and NusG, which bridge across the DNA binding channel to secure the DNA. Additionally, Spt5/NusG prevents structural changes in the RNAP that could otherwise lead to backtracking or

termination. Specifically, it is suggested that Spt4/5 ensures that the clamp remains in a firmly closed state. In eukaryotic complexes, a contact between RNA and Spt5 is formed by the linker domain between the KOW<sub>4</sub> and KOW<sub>5</sub> domains. In Archaea, however, only a single KOW domain is present (22) and the KOW domain does not contribute to the stimulatory effect of Spt4/5 on transcription elongation (20). Our reconstructions show that the overall architecture and organization of the TECs is conserved between Bacteria, Archaea and Eukaryotes, even though DNA contacts may vary among organisms and transcription systems. However, we do not observe direct contacts of the Spt5 KOW domain with the product RNA, even though this may depend on the scaffold and could change in a state in which the Spt5 KOW domain adopts a defined position.

Our structures are also well in line with results of smFRET studies that show that the conformational equilibrium between a contracted and intermediate state is shifted towards the intermediate conformation in the presence of Spt4/5.

Interestingly, while NusG, like archaeal Spt4/5, has anti-pausing activity in *E. coli*, no contacts between NusG and the NTS were reported so far for the TEC in *E. coli*. A multiple sequence alignment showed that the residues H12, R64, H65 and R67 are not highly conserved among NusG and Spt5 homologous. Notably, the histidine at position 12 is not even conserved among archaeal species and appears to be a unique residue in the NGN domain of Spt5 from *P. furiosus*. Mutating these residues led to a reduction of Spt4/5-dependent stimulation of RNAP's processivity (Figure 5C, D). We conclude that the Spt5 residues in the upstream-duplex interface are important for the function of Spt5 in *P. furiosus*, suggesting that archaeal Spt4/5 stimulates the processivity of the RNAP in a dual manner: (i) by its role as bridging factor that locks the DNA in the binding channel and (ii) by an interaction of the NGN domain with the upstream duplex to stabilize the transcription bubble. Similarly, the transcription initiation factor TFE stabilizes the upstream edge of the transcription bubble in initially transcribing complexes, demonstrating that a stabilization of the upstream edge is crucial to the integrity of transcription complexes (98). Most likely, such a stabilization by Spt5 or TFE is of even greater importance in hyperthermophilic Archaea as DNA melting and thus RNAP dissociation is favored by the high growth temperatures under which these organisms thrive. In this respect, TFE and Spt5 do not only physically but also functionally replace each other upon the factor-swapping mechanism in the transition from initiation to elongation (99). To which extent the interaction between the NTS and Spt5 is conserved across Archaea, however, remains to be determined by solving for example the elongation complex structures of well-known archaeal model systems like *Sulfolobus* or *Saccharolobus* species or from *Thermococcus kodakarensis*. For *T. kodakarensis* it was shown that Spt4/5 is crucial for the recruitment of the archaeal termination factor aCPSF1/FttA (38,39). Spt4/5 directly interacts with one of the aCPSF1 copies via residues F106, V131, V132 and I134. Moreover, aCPSF1 forms contacts with the stalk domain and the RNA. The KOW domain is not resolved in our structure indicating that this is a highly flexible domain that is stabilized by the interaction with aCPSF1. While only residue V131 is conserved between Spt5 from *P. furiosus* and *T. kodakarensis*, the overall fold of Spt5 homologues is highly similar, indicating that similar interactions can be formed in *P. furiosus*.

Taken together, we provide structural insights into archaeal transcription, which highlight the high degree of conformational flexibility of the clamp domain of the archaeal apo RNAP enzyme. We reveal that the apo RNAP-Spt4/5 complex adopts a super-contracted state, which hints to the existence of an inactivated RNAP pool in the cell. We furthermore provide the structure of an archaeal transcription elongation complex in the absence and presence of Spt4/5, revealing interactions of the Spt5-NGN domain with the upstream DNA, which are relevant for the function of Spt4/5 as an elongation factor.

## Data availability

The datasets generated and/or analyzed during the current study are available from the corresponding author on reasonable request. Structural data are deposited at the protein data bank: *Pfu* expanded apo RNA polymerase 8ORQ, EMD-17130; *Pfu* contracted RNA polymerase 8RBO, EMD-19033; *Pfu* RNA polymerase in complex with Spt4/5 8P2I, EMD-17366; Transcription Elongation complex 8CRO, EMD-16809; TEC bound to Spt4/5 8OKI, EMD-16929. The mass spectrometry proteomics data are deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD048268.

## Supplementary data

Supplementary Data are available at NAR Online.

## Acknowledgements

We would like to thank Sarah Schulz for single-molecule FRET measurements and analysis. We would furthermore like to thank Prof. Dr Ralph Witzgall and Prof. Dr Reinhard Rachel for JEM2100F access and training. Some cryo-EM data were collected at the cryo-EM facility of the University of Würzburg with support from Bettina Böttcher and Christian Kraft. We thank Henning Plikat for support with sample preparation for mass spectrometry analysis and Olexandr Dybkov for uploading the mass spectrometry datasets. With respect to single particle cryo-EM we would like to thank Rebecca Thompson for advice on sample preparation and Dr. Philipp Becker and Dr. Simona Pilotto for advice in model building. We furthermore thank David Pöllmann for the construction of the *Pyrococcus* strain used for RNAP purification, Martin Brehm for protein purification and crosslinking assistance. We thank Dipl. Ing. (FH) Thomas Hader and Simon Dechant for technical assistance to cultivate *P. furiosus* in large-scale bio-fermenters. Furthermore, we would like to thank all members of the transcription team at the Institute of Microbiology for fruitful discussions. The authors thank all past and present members of the structural biochemistry group within the RCB for their contribution, Norbert Eichner and Gerhard Lehmann for their invaluable IT support and Katharina Vogl and Mona Höcherl for technical assistance.

### Author contributions:

Christoph Engel: designed and supervised (cryo-)EM work, processing and data analysis, model-building and interpretation; contribution to funding, manuscript writing, discussion and editing, figure and movie generation.

Dina Grohmann: designed and supervised the research, biochemical data analysis, interpretation of *Pfu* RNAP struc-



tures, smFRET data analysis and interpretation, expression and purification of labelled RNAP and unlabelled Spt4/5 (*M. jannaschii* transcription system), securing funding, wrote the manuscript and created figures.

Felix Grünberger: cultivation of *P. furiosus* and initial mass spectrometry analysis of *Pfu* RNAP purified from cell mass, bioinformatical MS data analysis, bioinformatical analysis (multiple sequence alignment) and contribution to figures and material and methods section, discussion of the manuscript.

Winfried Hausner: creation of *P. furiosus* strain expressing tagged *Pfu* RNAP, *Pfu* RNAP preparation, initial experimental design and experiments for MS analysis of RNAPs, discussion of the manuscript.

Florian Heiß: cloning of *Pfu* Spt4/5 expression plasmids, support in (cryo-)EM sample preparation and processing.

Michael Pils: conducted and instructed (cryo-)EM work, sample preparation and optimization, processing and data analysis, model-building and interpretation; contribution to manuscript writing, discussion and editing, figure and movie generation.

Robert Reichelt: large scale and small-scale cultivation of *P. furiosus* wild-type and mutant strains, support in *Pfu* RNAP and Spt4/5 purification, support in transcription assay development, discussion of the manuscript.

Daniela Tarau: RNAP purification, RNAP negative staining, sample preparation for cryo-EM, establishing X-linking for TEC, cryo-EM data processing, model building, analysis and interpretation of *Pfu* RNAP structures, purification of Spt4/5 wt, cloning and purification of Spt4/5 mutants, transcription assays, preparation of *P. furiosus* for mass spectrometry analysis, Western Blot analysis, manuscript writing, creation of figures and movies

Sabine König, Henning Urlaub: LC-MS analysis.

## Funding

Deutsche Forschungsgemeinschaft [SFB960-TP7 to D.G. and SPP2141 to H.U.]; C.E. acknowledges funding by the Emmy-Noether-Programme [Deutsche Forschungsgemeinschaft, grant EN 1204/1-1]; Collaborative Research Center 960 [TP-A8]. The open access publication charge for this paper has been waived by Oxford University Press – NAR Editorial Board members are entitled to one free paper per year in recognition of their work on behalf of the journal.

## Conflict of interest statement

The authors state no conflict of interest. D.G. is co-founder of Microbify GmbH. However, there are no commercial interests by the company, or any financial support granted by Microbify GmbH.

## References

- Osman, S. and Cramer, P. (2020) Structural biology of RNA polymerase II transcription: 20 years on. *Annu. Rev. Cell Dev. Biol.*, **36**, 1–34.
- Griesenbeck, J., Tschochner, H. and Grohmann, D. (2017) Structure and function of RNA polymerases and the transcription machineries. In: Harris, J.R. and Marles-Wright, J. (eds.) *Macromolecular Protein Complexes, Subcellular Biochemistry*. Springer International Publishing, Cham, Vol. 83, pp. 225–270.
- Werner, F. and Grohmann, D. (2011) Evolution of multisubunit RNA polymerases in the three domains of life. *Nat. Rev. Micro.*, **9**, 85–98.
- Gehring, A.M., Walker, J.E. and Santangelo, T.J. (2016) Transcription regulation in archaea. *J. Bacteriol.*, **198**, 1906–1917.
- Chen, J., Boyaci, H. and Campbell, E.A. (2021) Diverse and unified mechanisms of transcription initiation in bacteria. *Nat. Rev. Micro.*, **19**, 95–109.
- Fernández-Tornero, C., Moreno-Morcillo, M., Rashid, U.J., Taylor, N.M.I., Ruiz, F.M., Gruene, T., Legrand, P., Steuerwald, U. and Müller, C.W. (2013) Crystal structure of the 14-subunit RNA polymerase I. *Nature*, **502**, 644–649.
- Engel, C., Sainsbury, S., Cheung, A.C., Kostrewa, D. and Cramer, P. (2013) RNA polymerase I structure and transcription regulation. *Nature*, **502**, 650–655.
- Hoffmann, N.A., Jakobi, A.J., Moreno-Morcillo, M., Glatt, S., Kosinski, J., Hagen, W.J.H., Sachse, C. and Müller, C.W. (2015) Molecular structures of unbound and transcribing RNA polymerase III. *Nature*, **528**, 231–236.
- Gietl, A., Holzmeister, P., Blombach, F., Schulz, S., Von Voithenberg, L.V., Lamb, D.C., Werner, F., Tinnefeld, P. and Grohmann, D. (2014) Eukaryotic and archaeal TBP and TFB/TF(II)B follow different promoter DNA bending pathways. *Nucleic Acids Res.*, **42**, 6219–6231.
- Werner, F. and Weinzierl, R.O.J. (2005) Direct modulation of RNA polymerase core functions by basal transcription factors. *Mol. Cell. Biol.*, **25**, 8344–8355.
- Smollett, K., Blombach, F., Reichelt, R., Thomm, M. and Werner, F. (2017) A global analysis of transcription reveals two modes of Spt4/5 recruitment to archaeal RNA polymerase. *Nat. Microbiol.*, **2**, 17021.
- Blombach, F., Fouqueau, T., Matelska, D., Smollett, K. and Werner, F. (2021) Promoter-proximal elongation regulates transcription in archaea. *Nat. Commun.*, **12**, 5524.
- Soppa, J. (1999) Transcription initiation in Archaea: facts, factors and future aspects. *Mol. Microbiol.*, **31**, 1295–1305.
- Ouhammouch, M., Dewhurst, R.E., Hausner, W., Thomm, M. and Geiduschek, E.P. (2003) Activation of archaeal transcription by recruitment of the TATA-binding protein. *Proc. Natl. Acad. Sci. U.S.A.*, **100**, 5097–5102.
- Bell, S.D., Kosa, P.L., Sigler, P.B. and Jackson, S.P. (1999) Orientation of the transcription preinitiation complex in Archaea. *Proc. Natl. Acad. Sci. U.S.A.*, **96**, 13662–13667.
- Nagy, J., Grohmann, D., Cheung, A.C.M., Schulz, S., Smollett, K., Werner, F. and Michaelis, J. (2015) Complete architecture of the archaeal RNA polymerase open complex from single-molecule FRET and NPS. *Nat. Commun.*, **6**, 6161.
- Grohmann, D., Nagy, J., Chakraborty, A., Klose, D., Fielden, D., Ebricht, R.H., Michaelis, J. and Werner, F. (2011) The initiation factor TFE and the elongation factor Spt4/5 compete for the RNAP clamp during transcription initiation and elongation. *Mol. Cell*, **43**, 263–274.
- Blombach, F., Salvadori, E., Fouqueau, T., Yan, J., Reimann, J., Sheppard, C., Smollett, K.L., Albers, S.V., Kay, C.W., Thalassinou, K., et al. (2015) Archaeal TFE $\alpha/\beta$  is a hybrid of TFIIE and the RNA polymerase III subcomplex hRPC62/39. *eLife*, **4**, e08378.
- Schulz, S., Gietl, A., Smollett, K., Tinnefeld, P., Werner, F. and Grohmann, D. (2016) TFE and Spt4/5 open and close the RNA polymerase clamp during the transcription cycle. *Proc. Natl. Acad. Sci. U.S.A.*, **113**, E1816–E1825.
- 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.
- 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: RNA

- polymerase-Spt4/5 complex architecture. *EMBO J.*, **30**, 1302–1310.
22. 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. U.S.A.*, **108**, 546–550.
  23. Werner, F. (2012) A nexus for gene expression—molecular mechanisms of Spt5 and NusG in the three domains of life. *J. Mol. Biol.*, **417**, 13–27.
  24. Decker, T.-M. (2021) Mechanisms of transcription elongation factor DSIF (Spt4–Spt5). *J. Mol. Biol.*, **433**, 166657.
  25. Sanders, T.J., Lammers, M., Marshall, C.J., Walker, J.E., Lynch, E.R. and Santangelo, T.J. (2019) TFS and Spt4/5 accelerate transcription through archaeal histone-based chromatin: transcription through archaeal chromatin. *Mol. Microbiol.*, **111**, 784–797.
  26. Landick, R. (2021) Transcriptional pausing as a mediator of bacterial gene regulation. *Annu. Rev. Microbiol.*, **75**, 291–314.
  27. Lawson, M.R., Ma, W., Bellecourt, M.J., Artsimovitch, I., Martin, A., Landick, R., Schulten, K. and Berger, J.M. (2018) Mechanism for the regulated control of bacterial transcription termination by a universal adaptor protein. *Mol. Cell*, **71**, 911–922.
  28. Wang, C., Molodtsov, V., Firlar, E., Kaelber, J.T., Blaha, G., Su, M. and Ebricht, R.H. (2020) Structural basis of transcription-translation coupling. *Science*, **369**, 1359–1365.
  29. Weixlbaumer, A., Grünberger, F., Werner, F. and Grohmann, D. (2021) Coupling of transcription and translation in Archaea: cues from the bacterial world. *Front. Microbiol.*, **12**, 661827.
  30. Webster, M.W., Takacs, M., Zhu, C., Vidmar, V., Eduljee, A., Abdelkareem, M. and Weixlbaumer, A. (2020) Structural basis of transcription-translation coupling and collision in bacteria. *Science*, **369**, 1355–1359.
  31. Blaha, G.M. and Wade, J.T. (2022) Transcription-translation coupling in bacteria. *Annu. Rev. Genet.*, **56**, 187–205.
  32. Yakhnin, A.V., FitzGerald, P.C., McIntosh, C., Yakhnin, H., Kireeva, M., Turek-Herman, J., Mandell, Z.F., Kashlev, M. and Babitzke, P. (2020) NusG controls transcription pausing and RNA polymerase translocation throughout the *Bacillus subtilis* genome. *Proc. Natl. Acad. Sci. U.S.A.*, **117**, 21628–21636.
  33. Nedialkov, Y., Svetlov, D., Belogurov, G.A. and Artsimovitch, I. (2018) Locking the nontemplate DNA to control transcription: rfaH remodels DNA in the transcription elongation complex. *Mol. Microbiol.*, **109**, 445–457.
  34. Kang, J.Y., Mooney, R.A., Nedialkov, Y., Saba, J., Mishanina, T.V., Artsimovitch, I., Landick, R. and Darst, S.A. (2018) Structural basis for transcript elongation control by NusG family universal regulators. *Cell*, **173**, 1650–1662.
  35. Vishwakarma, R.K., Qayyum, M.Z., Babitzke, P. and Murakami, K.S. (2023) Allosteric mechanism of transcription inhibition by NusG-dependent pausing of RNA polymerase. *Proc. Natl. Acad. Sci. U.S.A.*, **120**, e2218516120.
  36. Pilotto, S., Fouqueau, T., Lukoyanova, N., Sheppard, C., Lucas-Staat, S., Díaz-Santín, L.M., Matelska, D., Prangishvili, D., Cheung, A.C.M. and Werner, F. (2021) Structural basis of RNA polymerase inhibition by viral and host factors. *Nat. Commun.*, **12**, 5523.
  37. Engel, C., Neyer, S. and Cramer, P. (2018) Distinct Mechanisms of Transcription Initiation by RNA Polymerases I and II. *Annu. Rev. Biophys.*, **47**, 425–446.
  38. Sanders, T.J., Wenck, B.R., Selan, J.N., Barker, M.P., Trimmer, S.A., Walker, J.E. and Santangelo, T.J. (2020) FttA is a CPSF73 homologue that terminates transcription in Archaea. *Nat. Microbiol.*, **5**, 545–553.
  39. Marshall, C.J., Qayyum, M.Z., Walker, J.E., Murakami, K.S. and Santangelo, T.J. (2022) The structure and activities of the archaeal transcription termination factor Eta detail vulnerabilities of the transcription elongation complex. *Proc. Natl. Acad. Sci. U.S.A.*, **119**, e2207581119.
  40. Girbig, M., Misiaszek, A.D. and Müller, C.W. (2022) Structural insights into nuclear transcription by eukaryotic DNA-dependent RNA polymerases. *Nat. Rev. Mol. Cell Biol.*, **23**, 603–622.
  41. Wolberger, C. (2021) How structural biology transformed studies of transcription regulation. *J. Biol. Chem.*, **296**, 100741.
  42. Mohamed, A.A., Vazquez Nunez, R. and Vos, S.M. (2022) Structural advances in transcription elongation. *Curr. Opin. Struct. Biol.*, **75**, 102422.
  43. Nogales, E., Louder, R.K. and He, Y. (2017) Structural insights into the eukaryotic transcription initiation machinery. *Annu. Rev. Biophys.*, **46**, 59–83.
  44. Pilsl, M. and Engel, C. (2022) Structural studies of eukaryotic RNA polymerase I using cryo-electron microscopy. In: Entian, K.-D. (ed.) *Ribosome Biogenesis, Methods in Molecular Biology*. Springer US, NY, Vol. 2533, pp. 71–80.
  45. Kusser, A.G., Bertero, M.G., Naji, S., Becker, T., Thomm, M., Beckmann, R. and Cramer, P. (2008) Structure of an archaeal RNA polymerase. *J. Mol. Biol.*, **376**, 303–307.
  46. Jun, S.-H., Hirata, A., Kanai, T., Santangelo, T.J., Imanaka, T. and Murakami, K.S. (2014) The X-ray crystal structure of the euryarchaeal RNA polymerase in an open-clamp configuration. *Nat. Commun.*, **5**, 5132.
  47. Korkhin, Y., Unligil, U.M., Littlefield, O., Nelson, P.J., Stuart, D.I., Sigler, P.B., Bell, S.D. and Abrescia, N.G.A. (2009) Evolution of complex RNA polymerases: the complete archaeal RNA polymerase structure. *PLoS Biol.*, **7**, e1000102.
  48. Hirata, A., Klein, B.J. and Murakami, K.S. (2008) The X-ray crystal structure of RNA polymerase from Archaea. *Nature*, **451**, 851–854.
  49. Jun, S.-H., Hyun, J., Cha, J.S., Kim, H., Bartlett, M.S., Cho, H.-S. and Murakami, K.S. (2020) Direct binding of TFE $\alpha$  opens DNA binding cleft of RNA polymerase. *Nat. Commun.*, **11**, 6123.
  50. Sheppard, C., Blombach, F., Belsom, A., Schulz, S., Daviter, T., Smollett, K., Mahieu, E., Erdmann, S., Tinnefeld, P., Garrett, R., et al. (2016) Repression of RNA polymerase by the archaeo-viral regulator ORF145/RIP. *Nat. Commun.*, **7**, 13595.
  51. Wojtas, M.N., Mogni, M., Millet, O., Bell, S.D. and Abrescia, N.G.A. (2012) Structural and functional analyses of the interaction of archaeal RNA polymerase with DNA. *Nucleic Acids Res.*, **40**, 9941–9952.
  52. Fouqueau, T., Blombach, F., Hartman, R., Cheung, A.C.M., Young, M.J. and Werner, F. (2017) The transcript cleavage factor paralogue TFS4 is a potent RNA polymerase inhibitor. *Nat. Commun.*, **8**, 1914.
  53. Waegel, I., Schmid, G., Thumann, S., Thomm, M. and Hausner, W. (2010) Shuttle vector-based transformation system for *Pyrococcus furiosus*. *Appl. Environ. Microb.*, **76**, 3308–3313.
  54. Sommer, B., Waegel, I., Pöllmann, D., Seitz, T., Thomm, M., Sterner, R. and Hausner, W. (2014) Activation of a chimeric Rpb5/RpoH subunit using library selection. *PLoS One*, **9**, e87485.
  55. Fiala, G. and Stetter, K.O. (1986) *Pyrococcus furiosus* sp. nov. represents a novel genus of marine heterotrophic archaeobacteria growing optimally at 100°C. *Arch. Microbiol.*, **145**, 56–61.
  56. Gibson, D.G. (2011) Enzymatic assembly of overlapping DNA fragments. In *Methods in Enzymology*. Elsevier, Vol. 498, pp. 349–361.
  57. Garg, V.K., Avashthi, H., Tiwari, A., Jain, P.A. and Ramkete, P.W.R. (2016) MFPPi – Multi FASTA ProtParam Interface. *Bioinformatics*, **12**, 74–77.
  58. Pilsl, M., Heiss, F.B., Pöll, G., Höcherl, M., Milkereit, P. and Engel, C. (2022) Preparation of RNA polymerase complexes for their analysis by single-particle cryo-electron microscopy. In: Entian, K.-D. (ed.) *Ribosome Biogenesis, Methods in Molecular Biology*. Springer US, NY, Vol. 2533, pp. 81–96.
  59. Schorb, M., Haberbosch, I., Hagen, W.J.H., Schwab, Y. and Mastronarde, D.N. (2019) Software tools for automated transmission electron microscopy. *Nat. Methods*, **16**, 471–477.
  60. Alegria-Schaffer, A. (2014) General protein–protein cross-linking. In: *Methods in Enzymology*. Elsevier, Vol. 539, pp. 81–87.

61. Scheres, S.H.W. (2012) RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.*, **180**, 519–530.
62. Zheng, S.Q., Palovcak, E., Armache, J.-P., Verba, K.A., Cheng, Y. and Agard, D.A. (2017) MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods*, **14**, 331–332.
63. Bepler, T., Morin, A., Rapp, M., Brasch, J., Shapiro, L., Noble, A.J. and Berger, B. (2019) Positive-unlabeled convolutional neural networks for particle picking in cryo-electron micrographs. *Nat. Methods*, **16**, 1153–1160.
64. Tan, Y.Z., Baldwin, P.R., Davis, J.H., Williamson, J.R., Potter, C.S., Carragher, B. and Lyumkis, D. (2017) Addressing preferred specimen orientation in single-particle cryo-EM through tilting. *Nat. Methods*, **14**, 793–796.
65. Punjani, A., Rubinstein, J.L., Fleet, D.J. and Brubaker, M.A. (2017) cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods*, **14**, 290–296.
66. Wagner, T., Merino, F., Stabrin, M., Moriya, T., Antoni, C., Apelbaum, A., Hagel, P., Sitsel, O., Raisch, T., Prumbaum, D., et al. (2019) SPHIRE-crYOLO is a fast and accurate fully automated particle picker for cryo-EM. *Commun. Biol.*, **2**, 218.
67. Jakobi, A.J., Wilmanns, M. and Sachse, C. (2017) Model-based local density sharpening of cryo-EM maps. *eLife*, **6**, e27131.
68. (IUCr) Recent developments in the CCP-EM software suite.
69. Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F.T., de Beer, T.A.P., Rempfer, C., Bordoli, L., et al. (2018) SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.*, **46**, W296–W303.
70. Emsley, P. and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.*, **60**, 2126–2132.
71. Croll, T.I. (2018) ISOLDE: a physically realistic environment for model building into low-resolution electron-density maps. *Acta Crystallogr. D Struct. Biol.*, **74**, 519–530.
72. Pettersen, E.F., Goddard, T.D., Huang, C.C., Meng, E.C., Couch, G.S., Croll, T.I., Morris, J.H. and Ferrin, T.E. (2021) UCSF ChimeraX: structure visualization for researchers, educators, and developers. *Protein Sci.*, **30**, 70–82.
73. Adams, P.D., Afonine, P.V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.-W., Kapral, G.J., Grosse-Kunstleve, R.W., et al. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 213–221.
74. Atanassov, I. and Urlaub, H. (2013) Increased proteome coverage by combining PAGE and peptide isoelectric focusing: comparative study of gel-based separation approaches and peptide isoelectric focusing: comparative study of gel-based separation approaches. *Proteomics*, **13**, 2947–2955.
75. Roy, R., Hohng, S. and Ha, T. (2008) A practical guide to single-molecule FRET. *Nat. Methods*, **5**, 507–516.
76. Neyer, S., Kunz, M., Geiss, C., Hantsche, M., Hodorau, V.-V., Seybert, A., Engel, C., Scheffer, M.P., Cramer, P. and Frangakis, A.S. (2016) Structure of RNA polymerase I transcribing ribosomal DNA genes. *Nature*, **540**, 607–610.
77. Bernecky, C., Herzog, F., Baumeister, W., Plitzko, J.M. and Cramer, P. (2016) Structure of transcribing mammalian RNA polymerase II. *Nature*, **529**, 551–554.
78. Ehara, H., Yokoyama, T., Shigematsu, H., Yokoyama, S., Shirouzu, M. and Sekine, S. (2017) Structure of the complete elongation complex of RNA polymerase II with basal factors. *Science*, **357**, 921–924.
79. Bernecky, C., Plitzko, J.M. and Cramer, P. (2017) Structure of a transcribing RNA polymerase II–DSIF complex reveals a multidentate DNA–RNA clamp. *Nat. Struct. Mol. Biol.*, **24**, 809–815.
80. Wang, C., Molodtsov, V., Sanders, T.J., Marshall, C.J., Firlar, E., Kaelber, J.T., Santangelo, T.J. and Ebricht, R.H. (2023) Structural basis of archaeal FttA-dependent transcription termination. bioRxiv doi: <https://doi.org/10.1101/2023.08.09.552649>, 10 August 2023, preprint: not peer reviewed.
81. Weixlbaumer, A., Leon, K., Landick, R. and Darst, S.A. (2013) Structural basis of transcriptional pausing in bacteria. *Cell*, **152**, 431–441.
82. Landick, R. (2001) RNA polymerase clamps down. *Cell*, **105**, 567–570.
83. Lin, W., Das, K., Degen, D., Mazumder, A., Duchi, D., Wang, D., Ebricht, Y.W., Ebricht, R.Y., Sineva, E., Gigliotti, M., et al. (2018) Structural basis of transcription inhibition by fidaxomicin (Lipiamycin A3). *Mol. Cell*, **70**, 60–71.
84. Finn, R.D. (2000) Escherichia coli RNA polymerase core and holoenzyme structures. *EMBO J.*, **19**, 6833–6844.
85. Duchi, D., Mazumder, A., Malinen, A.M., Ebricht, R.H. and Kapanidis, A.N. (2018) The RNA polymerase clamp interconverts dynamically among three states and is stabilized in a partly closed state by ppGpp. *Nucleic Acids Res.*, **46**, 7284–7295.
86. Qayyum, M.Z., Molodtsov, V., Renda, A. and Murakami, K.S. (2021) Structural basis of RNA polymerase recycling by the Swi2/Snf2 family of ATPase RapA in Escherichia coli. *J. Biol. Chem.*, **297**, 101404.
87. Bellecourt, M.J., Ray-Soni, A., Harwig, A., Mooney, R.A. and Landick, R. (2019) RNA polymerase clamp movement aids dissociation from DNA but is not required for RNA release at intrinsic terminators. *J. Mol. Biol.*, **431**, 696–713.
88. Chakraborty, A., Wang, D., Ebricht, Y.W., Korlann, Y., Kortkhonjia, E., Kim, T., Chowdhury, S., Wigneshweraraj, S., Irschik, H., Jansen, R., et al. (2012) Opening and closing of the bacterial RNA polymerase clamp. *Science*, **337**, 591–595.
89. Fianu, I., Dienemann, C., Aibara, S., Schilbach, S. and Cramer, P. (2021) Cryo-EM structure of mammalian RNA polymerase II in complex with human RPAP2. *Commun. Biol.*, **4**, 606.
90. Aibara, S., Dienemann, C. and Cramer, P. (2021) Structure of an inactive RNA polymerase II dimer. *Nucleic Acids Res.*, **49**, 10747–10755.
91. Engel, C., Gubbey, T., Neyer, S., Sainsbury, S., Oberthuer, C., Baejen, C., Bernecky, C. and Cramer, P. (2017) Structural basis of RNA polymerase I transcription initiation. *Cell*, **169**, 120–131.
92. Tafur, L., Sadian, Y., Hoffmann, N.A., Jakobi, A.J., Wetzel, R., Hagen, W.J.H., Sachse, C. and Müller, C.W. (2016) Molecular structures of transcribing RNA polymerase I. *Mol. Cell*, **64**, 1135–1143.
93. Heiss, F.B., Daiß, J.L., Becker, P. and Engel, C. (2021) Conserved strategies of RNA polymerase I hibernation and activation. *Nat. Commun.*, **12**, 758.
94. Pilsl, M. and Engel, C. (2020) Structural basis of RNA polymerase I pre-initiation complex formation and promoter melting. *Nat. Commun.*, **11**, 1206.
95. Daiß, J.L., Pilsl, M., Straub, K., Bleckmann, A., Höcherl, M., Heiss, F.B., Abascal-Palacios, G., Ramsay, E.P., Tluczková, K., Mars, J.-C., et al. (2022) The human RNA polymerase I structure reveals an HMG-like docking domain specific to metazoans. *Life Sci. Alliance*, **5**, e202201568.
96. Liu, B. and Steitz, T.A. (2017) Structural insights into NusG regulating transcription elongation. *Nucleic Acids Res.*, **45**, 968–974.
97. Wang, Q., Daiß, J.L., Xu, Y. and Engel, C. (2022) Snapshots of RNA polymerase III in action – a mini review. *Gene*, **821**, 146282.
98. Blombach, F., Smollett, K.L., Grohmann, D. and Werner, F. (2016) Molecular mechanisms of transcription initiation—structure, function, and evolution of TFE/TFIIE-like factors and open complex formation. *J. Mol. Biol.*, **428**, 2592–2606.
99. Blombach, F., Daviter, T., Fielden, D., Grohmann, D., Smollett, K. and Werner, F. (2013) Archaeology of RNA polymerase: factor swapping during the transcription cycle. *Biochem. Soc. Trans.*, **41**, 362–367.