RNA Polymerase Recycling in *Bacillus subtilis*

Inaugural-Dissertation

to obtain the academic degree

Doctor rerum naturalium (Dr. rer. nat.)

submitted to the Department of Biology, Chemistry, Pharmacy

of Freie Universität Berlin

by

Haohong Pei

from China

2021

My doctoral studies were carried out since September 2017 under the supervision of Prof. Dr. Markus Wahl at the Department of structural biochemistry at the Free University of Berlin.

1. Reviewer: Prof. Dr. Markus Wahl
2. Reviewer: Prof. Dr. Florian Heyd

Date of defense: May 5th , 2021

Acknowledgements

I would like to thank my supervisor Prof. Dr. Markus Wahl for his outstanding supervision throughout my studies. I sincerely thank him for having given me the opportunity to join this lab, and work in an excellent scientific environment.

I would like to thank Prof. Dr. Florian Heyd for accepting the role of a second referee.

I would like to thank Dr. Tarek Hilal for cryoEM and negative stain EM analysis. I would like to thank Dr. Zhuo A. Chen for CLMS experiments and analysis. I would like to thank Dr. Bernhard Loll for helping me to solve the 3D structure of the RNAP-HelD-δ complex. I'd like to thank Prof. Dr. Irina Artsimovitch and Prof. Dr. Georgiy A. Belogurov for the helpful suggestions and talks. I thank Dr. Nelly Said for helping me regarding writing of my thesis. I thank Dr. Yong-Heng Huang for useful discussions about my projects. I would like to thank Yuan Gao for helping me with some biochemistry experiments. I would like to thank Dr. Karen Vester, Dr. Alexandra Bergfort, Lena Graß and Daniela Gjorgjevikj for corrections of my thesis.

I also thank all my lab colleagues for this memorable time, I thank them for their help and discussions.

I would like to thank the Chinese Scholarship Council and the Dahlem Research School for having given me financial support during my PhD study.

And last, I deeply have to thank my parents for being always there for me, their patience and great support.

Content

1 Summary

Transcription is the first step of gene expression, which is carried out by DNA-dependent RNA polymerases. RNA polymerases are considered to be well-regulated molecular machines that immediately initiate a new round of transcription after having been released from DNA and RNA during termination. However, RNA polymerases can become trapped in unproductive binary complexes with DNA or RNA. Polymerases trapped in this form endanger genome stability and lead to reduced pools of free polymerase. Moreover, RNA polymerases can enter dormant states. Recycling factors help retrieve RNA polymerases from trapped states, but their mechanisms remain elusive. Here we analyzed complexes of *Bacillus subtilis* RNA polymerase bound to a recycling ATPase, HelD, by cryo-electron microscopy, crosslinking / mass spectrometry and structure-informed biochemical analyses. HelD exhibits UvrD-like helicase domains from which a Gre-cleavage factor-like coiled-coil and a unique helical protrusion extend like two prongs. The coiled-coil inserts deep into the secondary channel of RNA polymerase, rearranges the active center and competes with bound RNA. The helical protrusion inserts into the primary channel, pushing the β and β' subunits apart and competing with downstream DNA. Insertion of the protrusion into the primary channel is aided by the intrinsically unstructured C-terminal region of the RNA polymerase δ subunit. The recovery of the polymerase is completed by ATP-mediated HelD release. We additionally observed a dimeric RNA polymerase-HelD complex, which suggests that HelD can also induce a dormant state at low ATP levels. Our results explain how HelD in collaboration with the δ subunit can rescue RNA polymerase entrapped on virtually any nucleic acid and suggest that HelD regulates transcriptional activity depending on the nutritional status of the cell.

2 Introduction

In all domains of life, transcription is the initial step in gene expression, where the genomic information saved in the DNA is transcribed into RNA by the multisubunit enzyme RNA Polymerase (RNAP). Three main types of RNA are obtained by transcription: messenger RNA (mRNA), which will be translated into amino acids; transfer RNA (tRNA), which transfers amino acids to ribosomes; ribosomal RNAs (rRNAs), which are involved in ribosome biogenesis and catalysis of protein synthesis. Transcription is performed in three steps: initiation, elongation and termination. Structural and functional studies have shed light on the molecular mechanism underlying transcription. However, cellular RNA Polymerases can become trapped on DNA or RNA, threatening genome stability and limiting free enzyme pools and it is still unknown how RNAP recycling into active states is achieved.

2.1 DNA-dependent RNA Polymerase

2.1.1 Organization of bacterial RNAP

RNA in all cellular organisms is synthesized by a complex molecular machine, the DNA-dependent RNA polymerase (RNAP). In its simplest bacterial form, the enzyme comprises at least five subunits, ß, ß', two α-subunits and ω, with a total molecular mass of around 400 kDa (Archambault, J. and Friesen, J.D. 1993). The essential core component of the bacterial RNAP (subunit composition $\alpha_2\beta\beta'$) is evolutionarily conserved from bacteria to humans.

The first high resolution RNAP structure was solved in 1999, a 3.3 Å crystal structure of *Thermus aquaticus* Core RNAP (Zhang, G.Y. et al., 1999). In the recent years, numerous RNAP structures have been solved, an *Escherichia coli (E. coli)* elongation complex (EC) architecture is shown in Figure 1.1 (Kang, J.Y. et al., 2017). The shape of RNAP is like a crab claw, with the two largest subunits ß and ß' forming the top and the bottom pincers. Between the pincers is the main channel, which contains the binding site for the nucleic acid and the active site (Figure 1.1). Two α subunits are located opposite to the pincers. Each α subunit consists of an N-terminal domain (NTD) and a C-terminal domain (CTD), in between is a flexible linker. The α-NTD contributes to RNAP assembly, serving as a scaffold to hold ß and ß' together, and the α-CTD (aa 250-329) is a DNA binding element and a major regulation target of transcription factors (Ito, K. et al., 1975; Jeon, Y.H. et al., 1997). Although the two α-CTDs are chemically and structurally identical, their interactions within RNAP and their functions differ. The ω subunit sits around the ß' subunit C-terminal tail. The ω subunit was reported to be involved in transcriptional control in response to nutrient

shifts, assisting the correct folding of the ß' subunit in assembly into the core RNAP (Mukherjee, K. et al., 1999; Vrentas, C.E. et al., 2005).

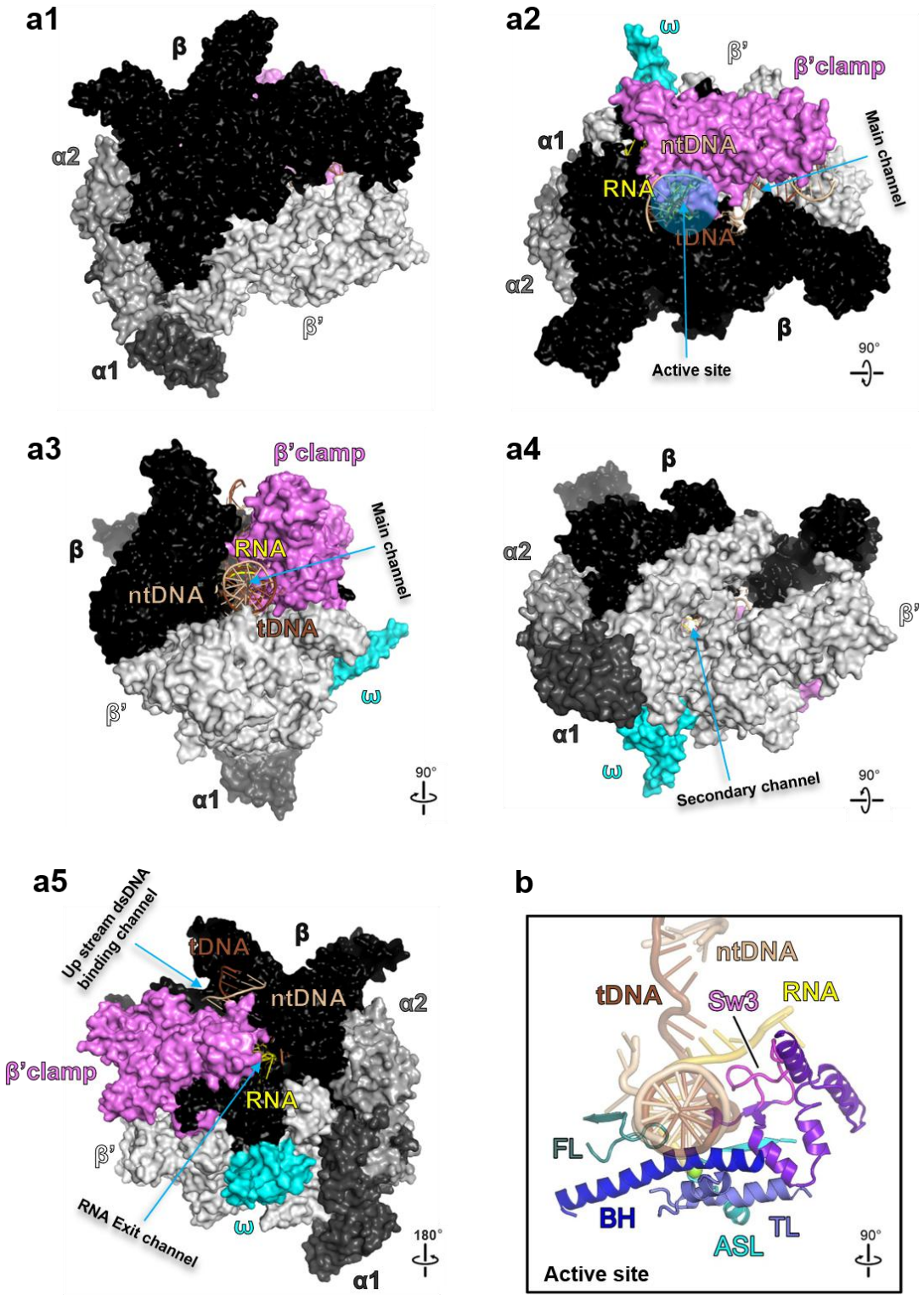


Figure 1.1 Structure of *E. coli* EC (Protein Data Bank (PDB) ID: 6ALF (Kang, J.Y. et al., 2017)). a1-a5, different side views of *E. coli* EC structure to show the organization of the complex and

each functional channel in different face of the complex. A transparent light blue dot region in a2 highlights the active site in the main channel. α1, dark gray; α2, gray; β, black; β', light gray; β' clamp, violet; ω, cyan; template (t) DNA, brown; non-template (nt) DNA, beige; RNA, gold. b, active site environment in an *E. coli* EC. FL (fork loop) (regions' name in RNAP are shown in Table 1.1), teal; C-term clamp (C-terminal clamp), purple; Sw3, magenta. β' elements: ASL, cyan; BH (bridge helix), blue; TL (trigger loop), slate blue; Mg^{2+} ion green sphere.

The two pincers are separated by a deep positively charged cleft, referred to as the main (primary) channel, which contains the active site at its back wall (the transparent light blue region in Figure 1.1 a2). The main channel harbors the binding site for the DNA-RNA hybrid and the downstream double-stranded DNA (dsDNA), which is held by the ß lobe (153-445; Table 1.1) and the ß' jaw (1149-1216, Table 1.1) elements (Cramer, P. et al., 2001; Ederth, J. et al., 2006; Borukhov, S. and Nudler, E. 2008).

The space within the main channel can be subdivided into three parts, with two minor channels branching off from the main channel to form the secondary channel and the RNA exit channel (Figure 1.1 a4, a5). The secondary channel provides the path for nucleoside triphosphate (NTP) substrates for RNA synthesis and for certain regulatory proteins and molecules (Zhang, G.Y. et al., 1999; Zenkin, N. and Yuzenkova, Y. 2015). The two pincers near the active center are connected via the ß' F-bridge α-helix (768-804) (Figure 1.1 b; Table 1.1), which forms one wall of the secondary channel. The RNA exit channel is built up of the upstream portions of ß (533-1342) and ß' pincer (clamp; aa 1-328) (Figure 1.1 a5), including the ß' "rudder" (308-328), "lid" (251-265) and the N-terminal "Zinc-finger" elements (ß' ZnF; 35-107), and the ß "fork loop" (533-549) and flexible "flap" (884-1046) (Table 1.1) (Vassylyev, D.G. et al., 2002; King, R.A. et al., 2004; Vassylyev, D.G. et al., 2007b).

Table 1.1 Regions of RNAP discussed in the text.

	E. coli	*B. subtilis*
β		
β1-lobe (protrusion)	31-139/456-512	33-128/412-468
β2-lobe	151-444	156-400
β SI1	226-350	281-367
β gate loop	359-388	229-258
β fork loop	533-599	489-557
β protrusion	450-507	406-463
β connector	814-839/1048-1065	773-798/907-924
β flap	830-1058	789-917
β flap tip	887-915	846-874
β flap tip arms	890-899/910-914	849-855/866-873
β SI2	938-1040	-
C-terminal β clamp	1233-1342	1038-1113
β switch 3	1247-1268	1052-1073 (not resolved)
β'		

β' ZBD	35-107	28-97
β' zipper	36-61	26-51
β' clamp	16-342/1318-1344	6-332/1132-1158
N-terminal β'clamp	132-190	122-176
β' lid	250-264	239-253
β' clamp helices	265-307	254-296
β' rudder	308-327	298-316
β' switch 1	1326-1327	1139-1140
β' switch 2	330-349	319-337 (not resolved)
β' dock	369-420	358-409
β' shelf	787-931	791-927
β' SI3	943-1130	-
β' jaw	1135-1317	952-1131
β' C-term	1318-1375	1132-1159
2° channel	480-790	473-794
β' bridge helix	768-850	777-853 (break 780-787)
Trigger loop	915-941/1130-1148	938-952

2.1.2 Additional subunits in Gram-positive bacterium *Bacillus subtilis*

Compared to RNAPs from Gram-negative bacteria, RNAPs from some Gram-positive bacteria, including *Bacillus subtilis*, contain two additional small subunits, ε and δ.

2.1.2.1 ε subunit

The small non-essential subunit ε had been identified as a ω-like protein (ω1) that copurifies with RNAP. However, the structure of epsilon revealed that it is not an omega subunit (Keller, A.N. et al., 2014). ε bears remarkable similarity to the Gp2 family of phage proteins involved in the inhibition of host cell transcription upon infection. Deletion of ε shows no phenotype and has no effect on the transcriptional profile of the cell (Keller, A.N. et al., 2014). Single-particle analysis suggests that it binds close to the downstream side of the DNA binding cleft (Figure 1.2) (Keller, A.N. et al., 2014). Because of the structural similarity of ε to Gp2 and the fact that they bind similar regions on RNA polymerase, it has been hypothesized that ε may serve a role in protection from phage infection (Keller, A.N. et al., 2014).

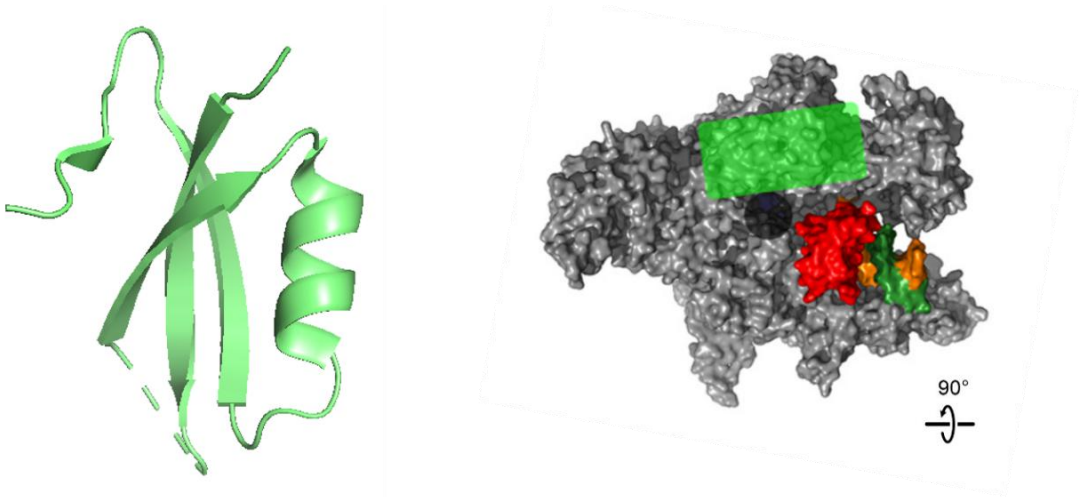


Figure 1.2 Structure of ε and its location on RNAP. left: Cartoon representation of the ε subunit structure (PDB 4NJC) at 2.3 Å resolution. right: Homology model of a *B. subtilis* RNAP elongation complex with the jaw region shown in red and the location of GFP marked in translucent green. The location of the secondary channel is marked as a dark gray circle. DNA coding (dark green) and noncoding (orange) strands as well as RNA transcript (blue) are also shown. Figure is adapted from Keller, A.N. et al., 2014, *Copyright © 2014, American Society for Microbiology*.

2.1.2.2 δ subunit

The 21.5 kDa δ subunit is a highly acidic protein and harbors two distinct regions, a 13 kDa amino-terminal domain (δ^{NTD}) with fairly uniform charge distribution and a glutamate and aspartate residue-rich carboxyl-terminal region (δ^{CTR}). The purified amino-terminal domain contains 32 % alpha-helix and 16 % beta-sheet, as judged by circular dichroism analysis (López de Saro, F.J. et al., 1995). In contrast, an 8.5 kDa tryptic fragment containing the C-terminal region is largely unstructured and highly charged. RNA polymerase purified from a *B. subtilis* mutant with an insertion in the δ gene (*rpoE::cat*) contains a truncated δ^{NTD}, indicating that the δ^{NTD} is stable *in vivo* and contains a RNAP core binding function (López de Saro, F.J. et al., 1995). The well-defined δ^{NTD} structure consists of four α-helices (helices I (Q8–K12), II(L16–H27), III(F33–L44), IV(G52–N63) and an antiparallel β-sheet, which is composed of three short β-strands (residues V31–P32, F68–A70, and T75–L78) at the top of a ‘‘twisted tripod’’ formed by helices II, III, and IV (Figure 1.3) (Motácková, V. et al., 2010; Papoušková, V. et al., 2013).

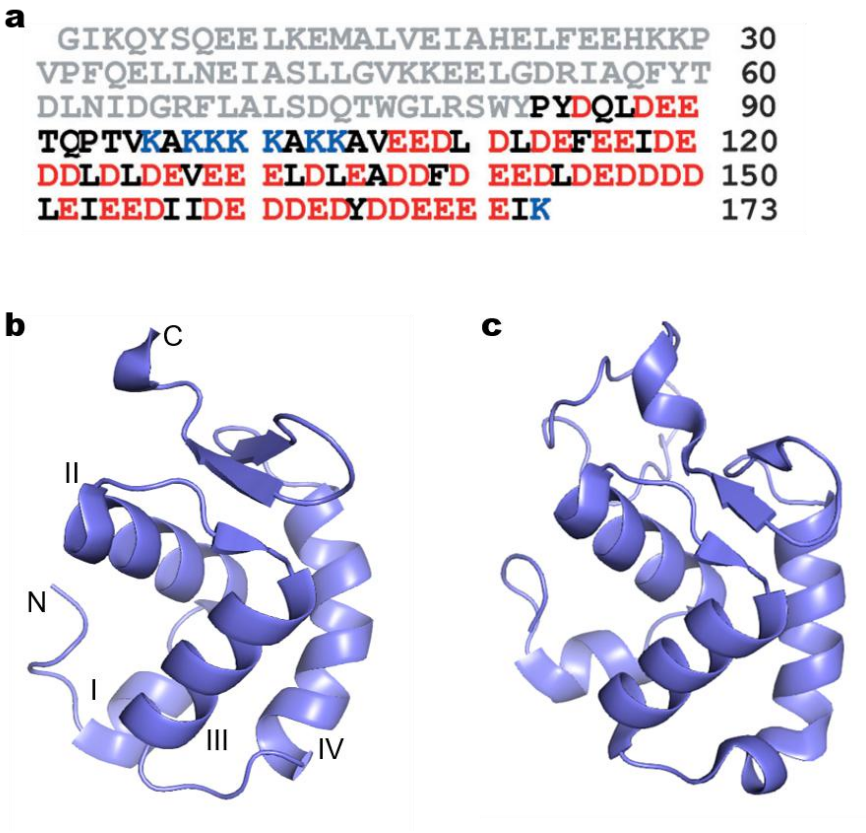


Figure 1.3 Structure of the δ subunit. a, Amino acid sequence of the δ subunit. The structured N-terminal domain is highlighted in gray; acidic and basic residues of the C-terminal domain are shown in red and blue, respectively. b, Representative structure of the N-terminal domain solved for the full-length δ construct (PDB ID: 2M4K). c, Structure determined for the truncated His-tagged N-terminal domain (PDB ID: 2KRC). (Papoušková, V. et al., 2013) *Copyright © 2013 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.*

δ can stimulate RNA synthesis in multiple rounds of *in vitro* transcription due to an increased rate of enzyme recycling, because δ can release stalled RNAP from the DNA template (Juang, Y.L. and Helmann, J.D. 1994). The displacement relies on δ^{CTR} since addition of δ, but not δ^{NTD}, displaces RNA bound to RNAP in a binary complex (López de Saro, F.J. et al., 1995). Although δ^{CTR} can displace nucleic acids from RNAP, this activity requires a large molar excess of protein and is relatively nonspecific in that both DNA and RNA are displaced (López de Saro, F.J. et al., 1995). The efficient displacement of RNA requires both binding of δ^{NTD} to RNAP and the negatively charged δ^{CTR}.

δ^{CTR} is strongly regulated by a short specific linker (the K-tract, K96-K104) between δ^{NTD} and δ^{CTR}, which is highly conserved in δ subunits from different bacteria. The presence of the K-tract has only a minor effect on local conformation, but greatly influences long-range contacts and overall shape of the molecule. The positive charged K-tract can interact with the negatively charged δ^{CTR}, which introduces a turn-like conformation of δ (Figure 1.4 top). Mutations of the K-tract to the opposite

charge (δ^{KE}) resulted in altered function of the mutant protein during transcription, leading to changes in the affinity of RNAP for DNA. This in turn negatively affected the competitive fitness of the cell (Kubáň, V. et al., 2019). The *in vivo* cross-linking studies indicated that δ^{CTR} is likely to enter the DNA binding channel of RNAP, where it can compete with incoming DNA (de Jong, L. et al., 2017). Most likely, binding of δ^{CTR} in the channel is transient. The dwelling time of δ^{CTR} outside the DNA binding channel may be prolonged by the interaction of δ^{CTR} with the K-tract, locking δ in a nonbinding conformation. Thus, the K-D/E motif ensures proper timing and kinetics of the movements of δ^{CTR} (Kubáň, V. et al., 2019).

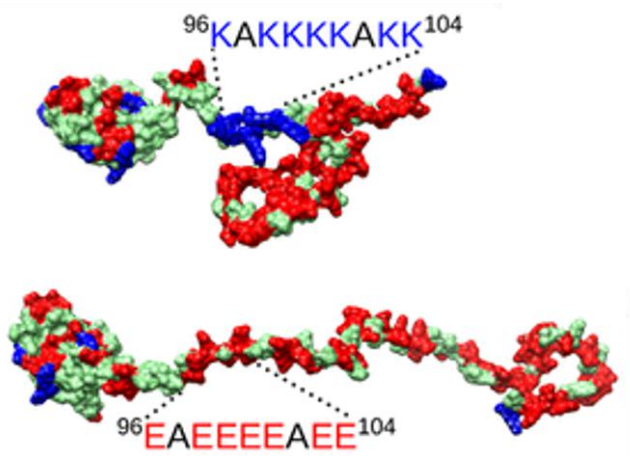


Figure 1.4 Comparison of full-length δ and δ^{KE}. top: δ^{CTR} can interact with K-tract to accommodate a turn-like structure. bottom: δ^{KE} cannot form this turn-like structure. (Kubáň, V. et al., 2019) *Copyright © 2019 American Chemical Society.*

Strains containing a deletion of the δ gene (*rpoE*) are viable and show no major alterations in gene expression, but have an altered morphology and are delayed in the exit from stationary phase (López de Saro, F.J. et al., 1999). Although not essential, δ is important for cell survival when facing a competing strain in a changing environment, as it is vital for the cell's ability to rapidly adapt and survive in nature (Rabatinová, A. et al., 2013).

2.1.3 Organization of eukaryotic RNAPs

In eukaryotic, transcription is carried out by three distinct nuclear RNA polymerases that transcribe different classes of genes. Protein-coding genes are transcribed by RNA polymerase II (Pol II) to produce mRNAs; ribosomal RNAs (rRNAs) and transfer RNAs (tRNAs) are transcripts of RNA polymerases I (Pol I) and III (Pol III). Pol I is responsible to transcribe the three largest species of rRNAs, which are termed 5.8S, 18S, and 28S according to their sedimentation rates during velocity centrifugation. Pol III transcribes the genes for tRNAs and for the smallest species of ribosomal RNA (5S rRNA). Some of the small RNAs that involved in protein transport and splicing

(scRNAs and snRNAs) are transcribed by Pol III as well, while others are transcribed by Pol II.

All three eukaryotic RNAP structures have been determined (Bernecky, C. et al., 2016: Ramsay, E.P. et al., 2020). Since RNAP dimerization will be discussed (in this thesis), which is a unique characteristic of Pol I, here we describe Pol I structure in detail. A crystal structure of *yeast* Pol I is shown in Figure 1.5 (Engel, C. et al., 2013). The *yeast* Pol I contains 14 subunits and has a total mass of 589 kDa. The Pol I core consists of 10 subunits: A190 and A135 are the largest subunits, which form the DNA binding cleft; five conserved subunits (Rpb5, Rpb6, Rpb8, Rpb10 and Rpb12) that present in all three Eukaryotic RNA polymerases; the AC40-AC19 heterodimer is homologous to Rpb3-Rpb11 in Pol II and conserved in Pol III; the last core subunit is A12.2, which participates in RNA cleavage (Figure 1.5 a, b) (Engel, C. et al., 2013). Except for the core, the A43-A14 heterodimer forms a stalk to provide the platform for transcription initiation factors and interacts with nascent RNA (Figure 1.5 a, b) (Kuhn, C.D. et al., 2007).

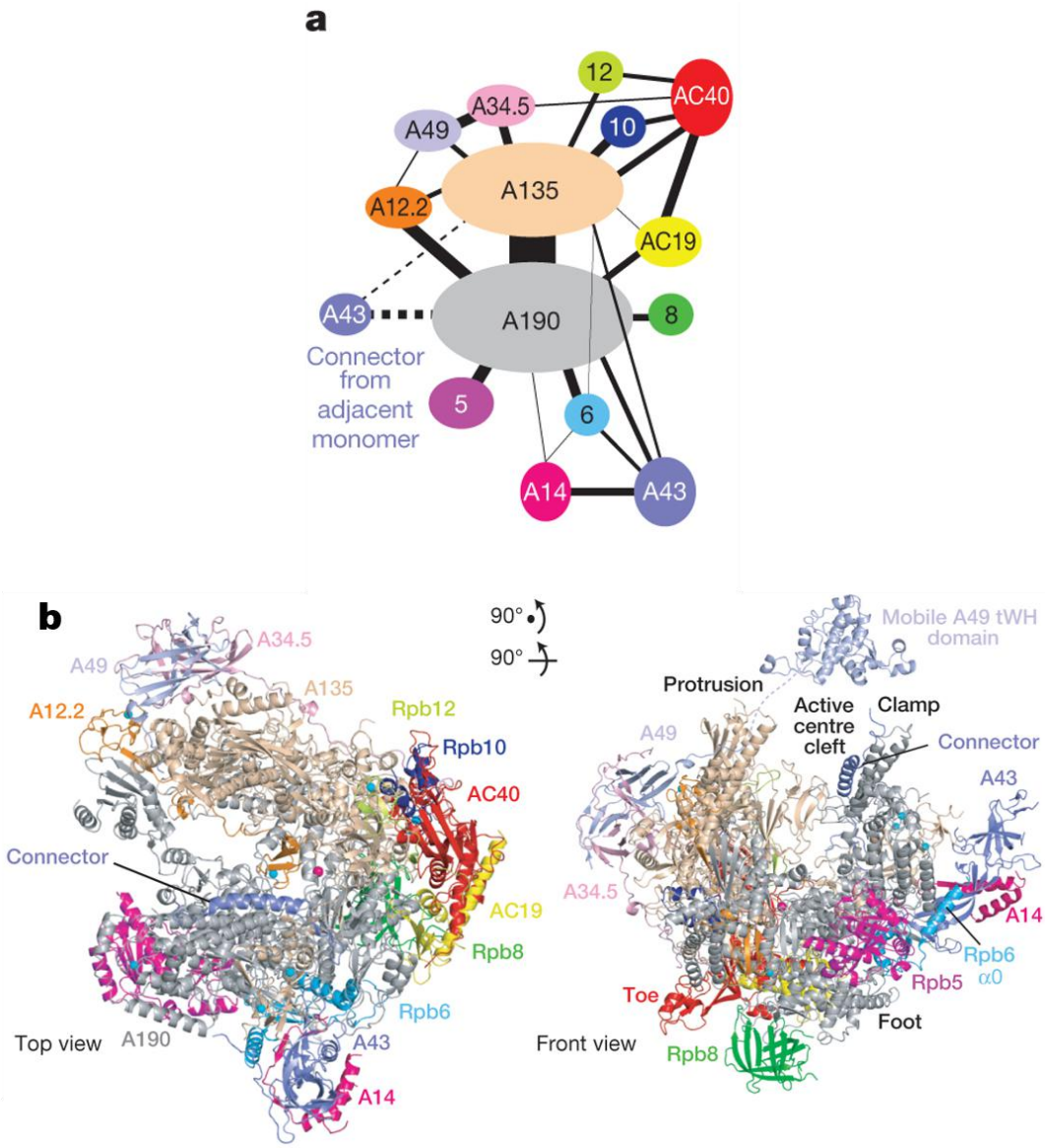


Figure 1.5 Crystal structure of Pol I. (PDB ID: 4C2M) a, Subunit interactions and colour key. The thickness of the connecting lines corresponds to the surface area buried in the corresponding interfaces. b, Top and side views of a ribbon representation of the Pol I structure. (Engel, C., Cramer, P. et al. 2013) *Copyright © 2013, Springer Nature.*

2.1.4 Pol I dimerization and hibernation

In vitro, Pol I was shown to form dimers (Milkereit, P., Schultz, P. and Tschochner, H. 1997). Pol I dimerization happens *in vivo* when cells are under stress conditions that leads to downregulation of rRNA synthesis (Torreira, E., Fernández-Tornero, C. et al. 2017). Nutrient starvation or the use of drugs blocking protein synthesis or ribosome biogenesis all trigger the formation of Pol I dimers (Neyer, S., Frangakis, A.S. et al. 2016; Pilsl, M., Schultz, P. et al. 2016). The dimerization is established by interactions between the A43-A14 heterodimer of one polymerase and the cleft region of another

polymerase (Figure 1.6 a, b) (Engel, C., Cramer, P. et al. 2013). The A43-A14 stalk is responsible for binding of transcription initiation factors, such as Rrn3, but this region is occluded in dimerized Pol I (Figure 1.6 c). Therefore, a dimerized Pol I is inactive. Hence, this dimerization represents a regulatory mechanism of Pol I under-stress conditions (Engel, C., Cramer, P. et al. 2013).

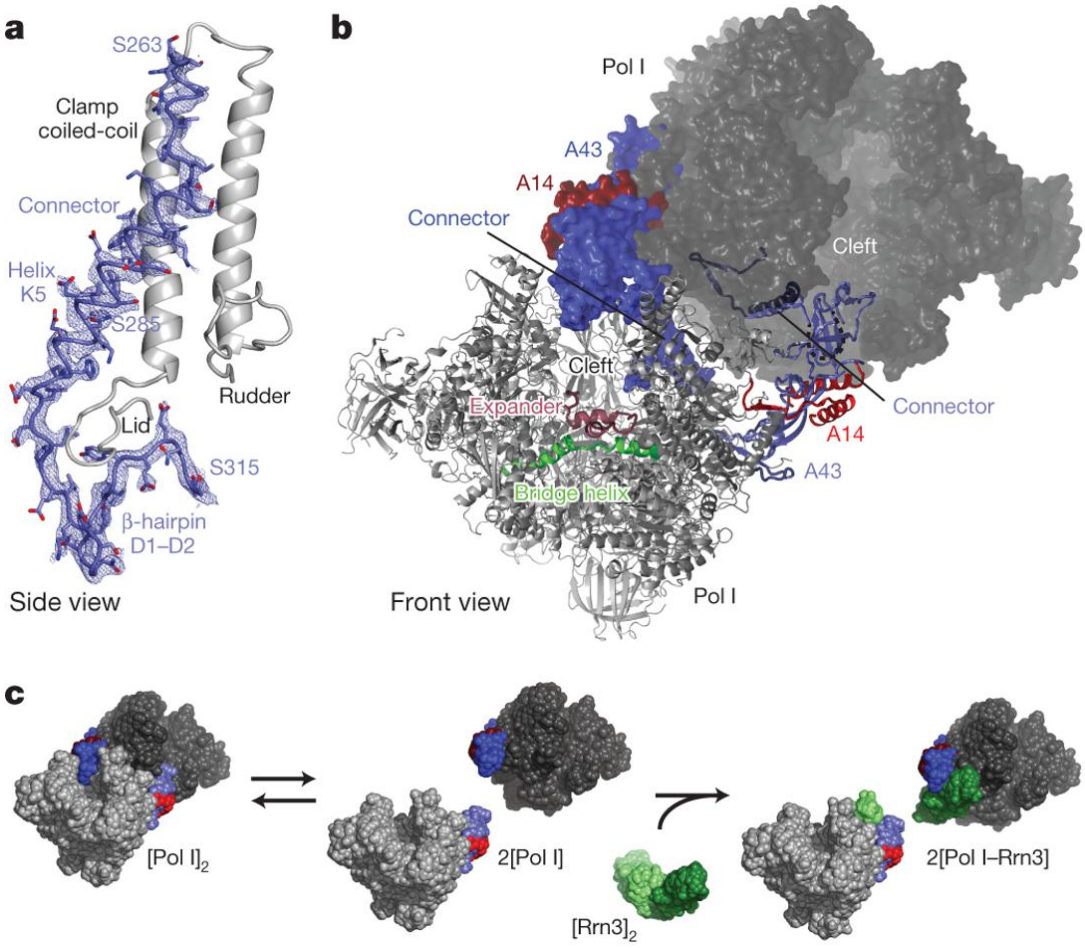


Figure 1.6 Dimerization of Pol I and model for Pol I initiation regulation. a, Structure of the A43 connector and corresponding $2F_o - F_c$ electron density (blue mesh, contoured at 1σ) and its binding to the coiled-coil at the inner side of the clamp (silver). b, Structure of the Pol I dimer and its stabilization by the A14–A43 heterodimer. One polymerase is shown as a cartoon model, the other in surface representation. The connector (blue) invades the cleft of a neighboring polymerase. c, Model for Pol I initiation regulation. The binding site for the initiation factor Rrn3 (green) is occluded in the inactive Pol I dimer (silver), A14 and A43 highlighted in red and blue, respectively. Rrn3 is released from inactive Rrn3 dimers, resulting in active Pol-I–Rrn3 complexes. (Engel, C. et al., 2013) *Copyright © 2013, Springer Nature.*

Pol I dimerization can be seen as a hibernating mechanism under stress conditions. Pol I hibernation might protect the enzyme from degradation and allow fast reactivation when favourable growth conditions are restored. A similar mode of hibernation by dimerization has also been found for bacterial ribosomes (Beckert, B. et al., 2017; Khusainov, I. et al., 2017). However, while the formation of ribosome

homodimers relies on external factors that bind prior to dimerization, Pol I dimerization does not require external factors (Engel, C. et al., 2013; Fernández-Tornero, C. et al., 2013). Control of the Pol I monomer-dimer transition relies on yet unknown regulatory mechanisms. In addition to dimerization, hibernation implies cleft expansion and ordering of the DNA-mimicking loop inside the cleft. In case of Pol II and bacterial RNA polymerase, certain RNAs and proteins can block the enzyme by binding inside the cleft (Chen, J. et al., 2017; Kettenberger, H. et al., 2006). The DNA-mimicking loop within the expander could have a protective function in the Pol I hibernating state, by hampering the binding of macromolecules that could compromise enzyme reactivation (Fernández-Tornero, C. 2018).

2.2 Transcription cycle

In general, transcription cycling consists of three steps: initiation, elongation and termination (Figure 1.7). Transcription initiation is regulated by initiation factors called σ, and the RNAP core enzyme, composed of $\alpha_2\beta\beta'\omega$, will form a holoenzyme with one of these σ factors. σ factors recognize specific sequences on the DNA template, called promoter region, and, together with the holoenzyme, initiates transcription. After promoter escape, elongation factors including NusA, NusG will replace σ to form an elongation complex (EC) (Gill, S.C. et al., 1991; Mooney, R.A. et al., 2009). The Nus transcription-factors were originally identified as part of the *E. coli* phage λ N-protein-controlled antitermination system; hence, they were termed N-utilization substances (Friedman, D.I. and Baron, L.S. 1974). NusA and NusG are the most conserved proteins among the Nus transcription factors (Nudler, E. and Gottesman, M.E. 2002; Werner, F. 2012). NusA interacts with RNAP and nascent RNA, which enhances RNA hairpin-induced transcription pauses, leading to increase transcription termination (Landick, R. and Yanofisky, C. 1987; Gusarov, I. and Nudler, E. 1999). NusG accelerates the rate of RNA chain elongation (Herbert, K.M. et al., 2010). Elongation complex can transform into an arrested state, in which the RNAP-DNA-RNA ternary complex neither elongates nor dissociates. In this case, transcript cleavage factors GreA/B stimulate RNA cleavage within ternary complexes to resume transcription elongation (Borukhov, S. et al., 1993). The ensuing elongation complex moves along the DNA template and synthesizes RNA until a termination signal is reached. When the EC encounters a termination signal, the complex will be destabilized and the transcript will be released in a matter of seconds, hence the transcription is terminated (Wang, D. et al., 2006; Santangelo, T.J. and Artsimovitch, I. 2011).

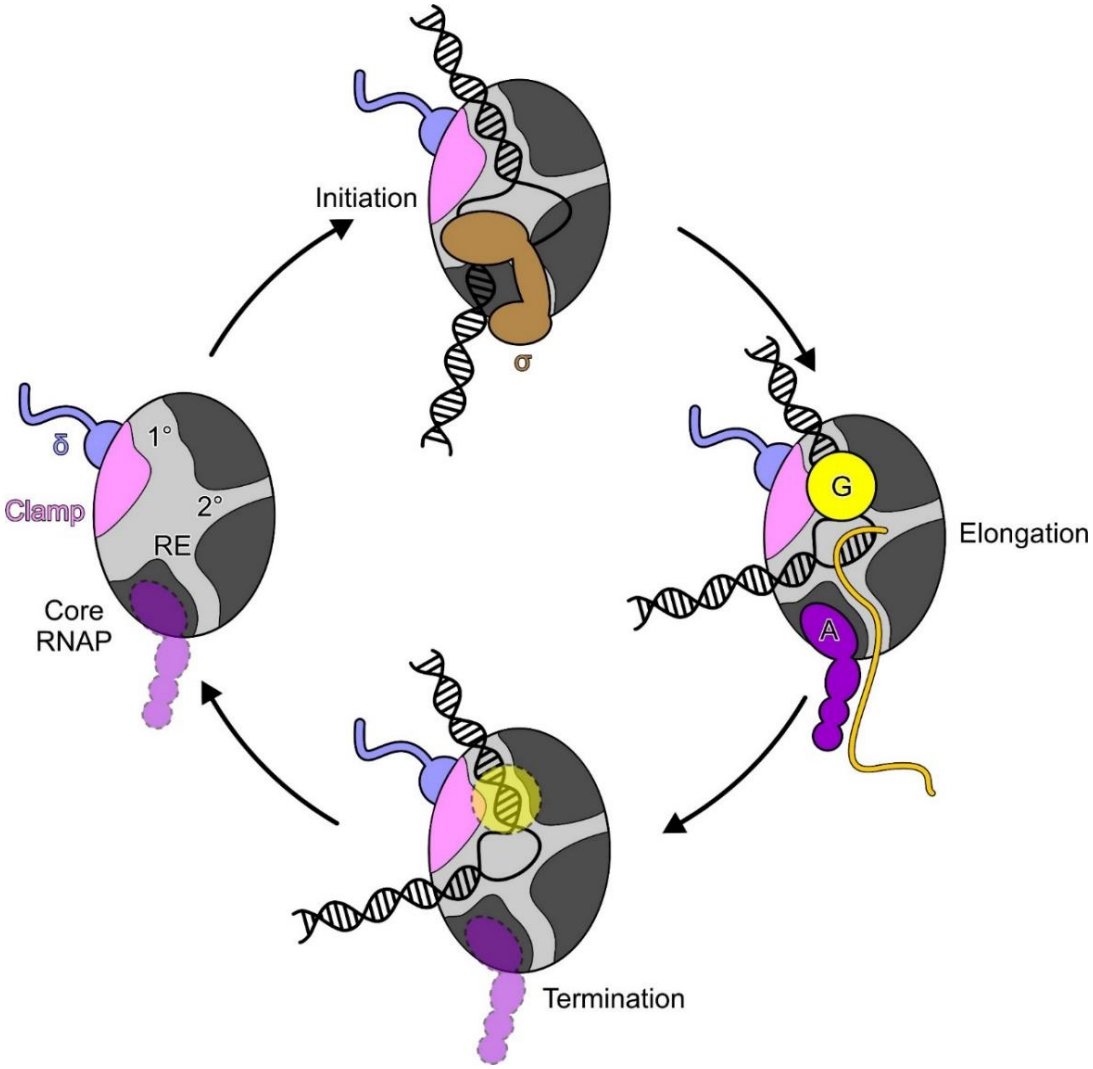


Figure 1.7 transcription cycle model. 1°/2°, main/secondary channels; RE, RNA exit tunnel; A/G, general elongation factors NusA/NusG. NusG binds across the active center cleft, while NusA binds next to the RNA exit channel at the β FT (ß “fork loop”). Semitransparent icons with dashed lines indicate that the respective factor may be released at the respective step.

2.2.1 Transcription initiation factors in *Bacillus subtilis*

σ factors are transcription initiation factors, which bind to core RNAP to recognize specific promoter sequences and thus to transcribe the genes. *Bacillus subtilis* has 17 different σ factors, which are synthesized and activated at various times during development or after changes in environmental conditions. σ^A is the principal σ factor present in vegetatively growing *Bacillus subtilis* (Shorenstein, R.G. and Losick, R. 1973). The σ^A protein has an apparent molecular mass of 55 kDa. The addition of purified σ^A to core RNAP stimulates *in vitro* transcription from several *Bacillus subtilis* DNAs (Chang, B.Y. and Doi, R.H. 1990; Tjian, R. et al., 1977). σ^B is an alternative factor with an apparent molecular mass of 37 kDa that is activated in response to environmental stress or energy depletion (Boylan, S.A. et al., 1993). The general

stress regulon under the control of σ^B provides the cell with multiple stress resistances (Petersohn, A. et al., 2001).

2.2.2 Transcription termination in *Bacillus subtilis*

In prokaryotes, termination can be divided in two classes: intrinsic termination and factor-dependent termination. The factor-dependent termination includes Rho(ρ)-dependent termination and Mfd-dependent termination. The intrinsic termination signal is encoded in DNA, which consist of a GC-rich inverted repeat, followed by a stretch of 4-9 T nucleotides. Because of this special construction, the 3' terminus of nascent RNA can form a hairpin followed by a U-stretch (d'Aubenton Carafa, Y. et al., 1990; Gusarov, I. and Nudler, E. 1999). The EC termination is carried out by two mayor steps. Firstly, the formation of a short stem-loop in the nascent RNA is the initial force to start the termination (Epshtein, V. et al., 2007). During the formation of the stem-loop, the main channel of RNAP is invaded by the hairpin, initiating the melting of the DNA:RNA hybrid. Secondly, the further elongated hairpin reaches the active site, inducing conformational changes. Thus, the nucleotide addition is blocked and followed by conformational changes in other parts of RNAP. Due to these changes, all protein-nucleic acid interactions stabilizing the transcription elongation complex have been disrupted and thus nascent RNA dissociated.

In Gram-negative bacteria, eg. *E. coli*, ρ-dependent termination plays an important role for transcription termination. However, in the Gram-positive bacteria *B. subtilis*, ρ protein is dispensable (Ingham C.J. et al., 1999), and intrinsic termination is dominant in *B. subtilis* (de Hoon, M.J.L. et al., 2005). The only known case of ρ-dependent termination in *B. subtilis* is the *rho* gene itself. The absence of *rho* gene only caused mild phenotypes (Ingham C.J. et al., 1999).

2.2.3 Transcription-translation uncoupling in *Bacillus subtilis*

Gene expression is performed in two steps: transcription and translation. Firstly, RNAP using DNA as the template to transcribe the mRNA. Secondly, the ribosome utilizes mRNA as the template to synthesis proteins. In eukaryotes, these two steps are spatially separated. While transcription happens in the cell nucleus, translation occurs in the cytoplasm afterwards. However, it was found transcription and translation are coupled in *E. coli*, which indicated that transcription and translation are spatially close (Miller Jr., O.L. et al., 1970) and ribosome and RNAP are connected via transcription factor NusG (Burmann, B.M. et al., 2010; Saxena, S. et al., 2018). Recent published works determined a series of cryoEM structures of *E. coli*

RNAP-ribosome supramolecular complexes, which shows that ribosome and RNAP can either be connected via NusG or linked by mRNA (Webster, M.W. et al, 2020; Wang, C.Y. et al., 2020). The question arises, whether transcription-translation coupling is indeed a fundamental characteristic in all bacteria species other than *E. coli*. A recent study reported that transcription-translation is uncoupled in *B. subtilis*, which often has markedly divergent regulatory mechanisms compared to *E. coli* (Johnson, G.E. et al., 2020). In contrast to translation-controlled termination of transcription in *E. coli*, *B. subtilis* RNAPs is insensitive to translation, since it was reported that the intrinsic terminator is effective even when the entire terminator hairpin is translated (Johnson, G.E. et al., 2020). And in *B. subtilis*, ρ is not utilized to remove the aberrant nonsense mRNAs from RNAP, a mechanism conducted in *E. coli* (Johnson, G.E. et al., 2020).

2.3 RNAP recycling

RNAPs are viewed as well-tuned engines that promptly re-initiate a new round of transcription after termination. However, RNAP can linger on DNA after RNA release (Kang, W. et al., 2020; Bellecourt, M.J. et al., 2019; Harden, T.T. et al., 2020), roadblocking replisomes which trigger double-stranded DNA breaks (Washburn, R.S. and Gottesman, M.E. 2011) and giving rise to aberrant antisense transcripts (Harden, T.T. et al., 2020). RNAP can also form binary complexes with RNA (López de Saro, F.J. et al., 1995; Altmann, C.R. et al., 1994), either through *de novo* association with stable RNAs, such as tRNAs and 6S RNA (Busby, S., Spassky, A. and Buc, H. 1981; Wassarman, K.M. and Saecker, R.M. 2006), or in the course of hairpin-induced termination (Epshtein, V. et al., 2007). While some RNA binary complexes serve as RNAP storage depots and can be reactivated when nutrients become available (Wassarman, K.M. and Saecker, R.M. 2006), others may sequester unproductive RNAP (Travers, A. 1976). Thus, RNAP has to be recycled to promote another round of transcription.

2.3.1 RapA in *Escherichia coli* RNAP recycling

The ATPase RapA, a 110 kDa protein, was identified in 1998 by co-purification with RNAP. The ATPase activity of RapA is stimulated by binding to RNAP. RapA has a putative helicase motif, but showed no detectable helicase activity in helicase assays using several different substrates (Sukhodolets, M.V. and Jin, D.J. 1998). RapA has an amplification effect on transcription by stimulating RNAP recycling, resulting in effective multiple-round transcription (Sukhodolets, M.V. et al., 2001). The crystal structure of RapA (PDB ID: 3DMQ) reveals that the protein contains seven domains: an N-terminal domain (NTD), two RecA-like domains 1A and 2A, two Swi2/Snf2-specific domains 1B and 2B, a Spacer domain, and a C-terminal domain (Figure 1.8 A) (Shaw, G. et al., 2008). Domains 1A, 2A, 1B, and 2B form the

Swi2/Snf2-specific ATPase module of the protein.

The ATPase activity of RapA is essential for its function as a transcriptional activator, but it is inhibited by its NTD in the isolated RapA. ATPase activity is stimulated by binding to RNAP, which releases the autoinhibitory effect of the RapA-NTD, supporting RNAP recycling (Kakar, S. et al., 2015).

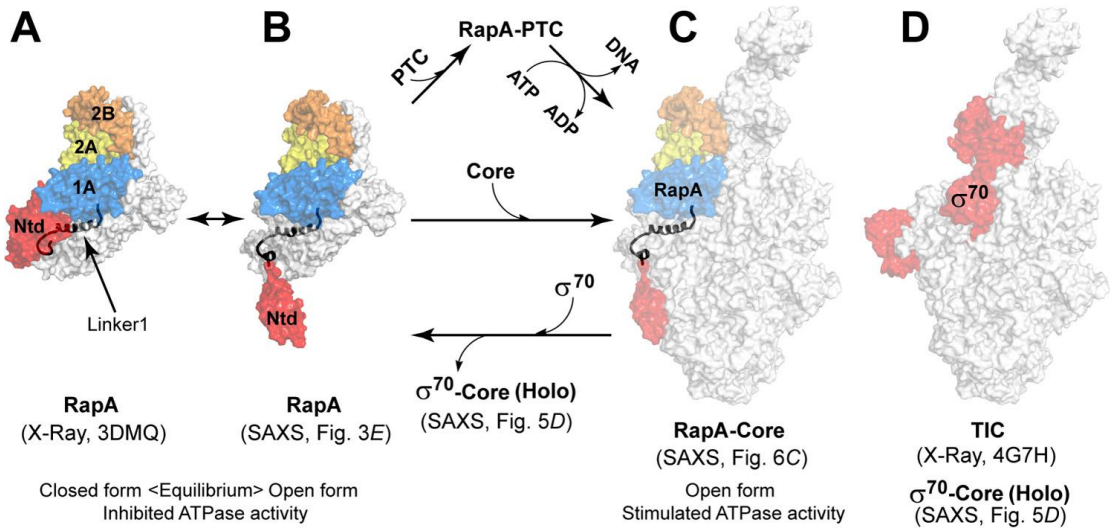


Figure 1.8 The functional cycle of RapA. A, the closed form of *E. coli* RapA (EcRapA) (PDB ID 3DMQ) shows that the NTD is packed against the ATPase module of RapA. B, the open form of EcRapA according to SAXS shows that the NTD is disengaged from the ATPase module. C, the SAXS structure of EcRapA-TtCore (*T. thermophilus* RNAP core) shows that the open form of EcRapA is consistent with its binding mode to the TtCore. Proteins are shown as transparent molecular surfaces, except for Linker1, which is shown as a ribbon diagram; proteins are colored in white, except that the NTD, 1A, 2A, 2B, and Linker1 of EcRapA colored in red, cyan, yellow, orange, and black, respectively. D, The crystal structure of TtHolo (*T. thermophilus* RNAP-σ holoenzyme) from the transcription initiation complex (TtTIC; PDB ID 4G7H) is illustrated as a molecular surface in white, with σ^{70} highlighted in red. The promoter DNA is omitted from the structure. (Kakar, S. et al., 2015) *Copyright © 2015 by The American Society for Biochemistry and Molecular Biology, Inc.*

2.3.2 Mfd releases RNAP from the DNA template in transcription-repair coupling

Nucleotide excision repair (NER) helps to ameliorate the lethal and mutagenic consequences of DNA damage by removing helix-distorting lesions from cellular genomes (Hanawalt, P.C. and Haynes, R.H. 1967). In 1989, transcription-repair coupling (TRC) was first described in *E. coli*, as it was found that the NER in the transcription induced DNA strand is significantly faster than in the non-transcribed strand (Mellon, I. and Hanawalt, P.C. 1989). However, NER can be inhibited by template DNA blocked RNAP. A transcription-repair coupling factor called Mfd was

found to remove the blocked RNAP and deliver the excision repair enzymes (UvrABC), to the lesion, thereby mediating more rapid repair of the transcription-blocking lesion and coupling DNA repair with transcription (Figure 1.9) (Selby, C.P. and Sancar, A. 1990; Selby, C.P. and Sancar, A. 1991; Svoboda, D.L. et al., 1993; Selby, C.P. and Sancar, A. 1993). At the same time, nascent RNA can be released via Mfd caused forward translocation of RNAP by using the ATP-dependent translocase activity of Mfd (Roberts, J. and Park, J.S. 2004). Therefore, Mfd is also considered as a termination factor. Structural and functional analyses of Mfd protein revealed its helicase motifs to be responsible for ATP hydrolysis and DNA binding, and identified the regions that interact with RNAP and the repair enzyme UvrA (Deaconescu, A.M. et al., 2006).

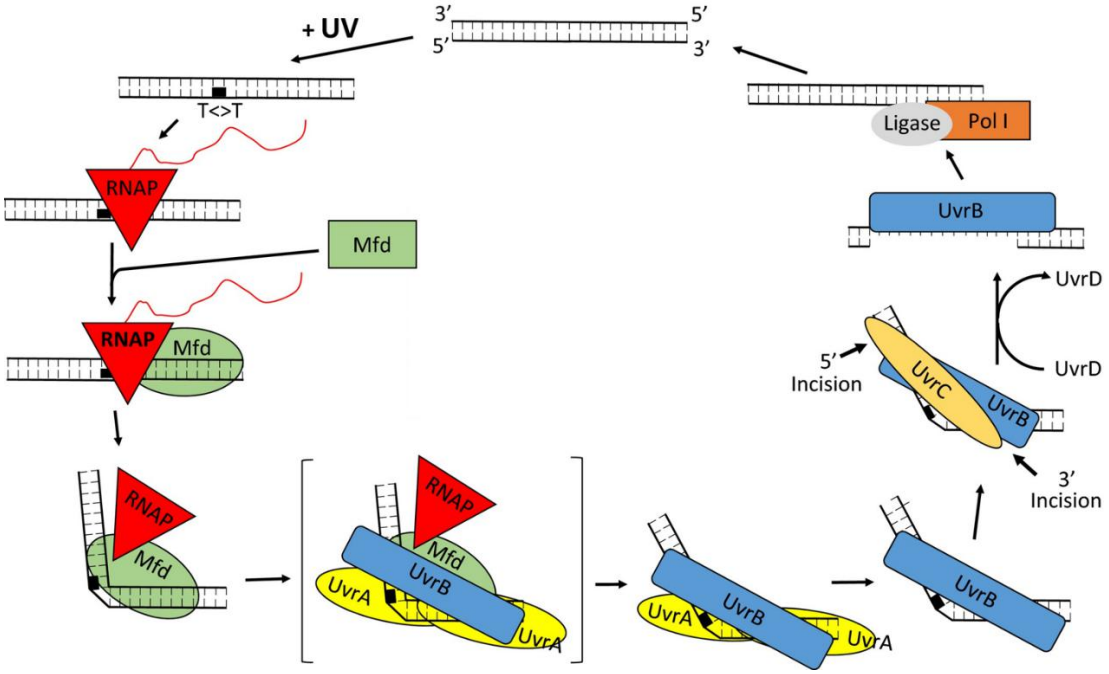


Figure 1.9 Model for transcription–repair coupling (TRC) in *E. coli*. Mfd in its "closed", repressed state (rectangle), exhibits an altered, "open," active state (oval) after binding to RNAP blocked by a template strand lesion. The active state Mfd exhibits higher ATPase activity, binds with higher affinity to DNA and exhibits translocase activity. Upon binding to RNAP, a DNA structure with a 90-degree bend forms, in which the DNA is partially wrapped around Mfd. Also, RNAP and RNA are dissociated from the template by Mfd, and RNAP remains tethered to Mfd. In this conformation, Mfd binds strongly to the 2UvrA–1UvrB damage recognition complex. Binding of UvrA–UvrB to Mfd at this stage results in dissociation of Mfd and RNAP, and UvrA–UvrB are delivered to the damage site. The figure shows the arriving Uvr proteins and the departing Mfd and RNAP proteins in brackets to represent a transient intermediate. The following steps of repair are the same as in transcription-independent repair. (Selby, C.P. 2017) *Copyright © 2016 The American Society of Photobiology.*

2.3.3 Rho-mediated transcription complex disassembly

ρ-dependent termination depends on the termination factor ρ (Ciampi, M.S. 2006), which is a hexameric ring-shaped RecA-family RNA translocase, existing in open and closed configuration responsible for loading onto RNA and translocation, respectively (Thomsen, N.D. et al., 2016). ρ-dependent termination signals generally consist of a over 30 nt long pyrimidine-rich RNA element which is called ρ utilization (*rut*) site. It was reported that ρ already associates with the early transcription complex and that ρ starts termination actions once it encounters a certain signal sequence in the nascent RNA (Mooney, R.A. et al., 2009; Epshtein, V. *et al.,* 2010).

A very recent work from our group, by determining a series of ρ associated NusA/NusG-EC complex cryoEM structures, proposed a ρ-mediated EC disassembly pathway, in which accompanying RNAP and transcription factors NusA and NusG play key roles (Figure 1.10) (Said, N. *et al.*, 2020). The open-ring shape ρ traffic on proximal upstream DNA (prox. uDNA), and recruited by an NusA-EC. ρ may passively tramslocate on an NusA-EC in an open conformation. NusG entry into the complex leads to the repositioning of NusA, capturing the prox. uDNA, which is locally unwound by ρ_1 at primary DNA binding site. ρ interacts with NusA, $NusG^{NTD}$ and several regions of RNAP. Until the complex encounters a pause site, NusA wedges into the ρ ring, initially sequestering RNA. While distal upstream DNA is redirected over the Zinc-binding domain of RNAP, NusA turns underneath ρ_6, which in turn captures the RNA. After the displacement of NusG and the opening of RNAP clamp, RNAP loses the binding of RNA:DNA hybrid and transcription terminates.

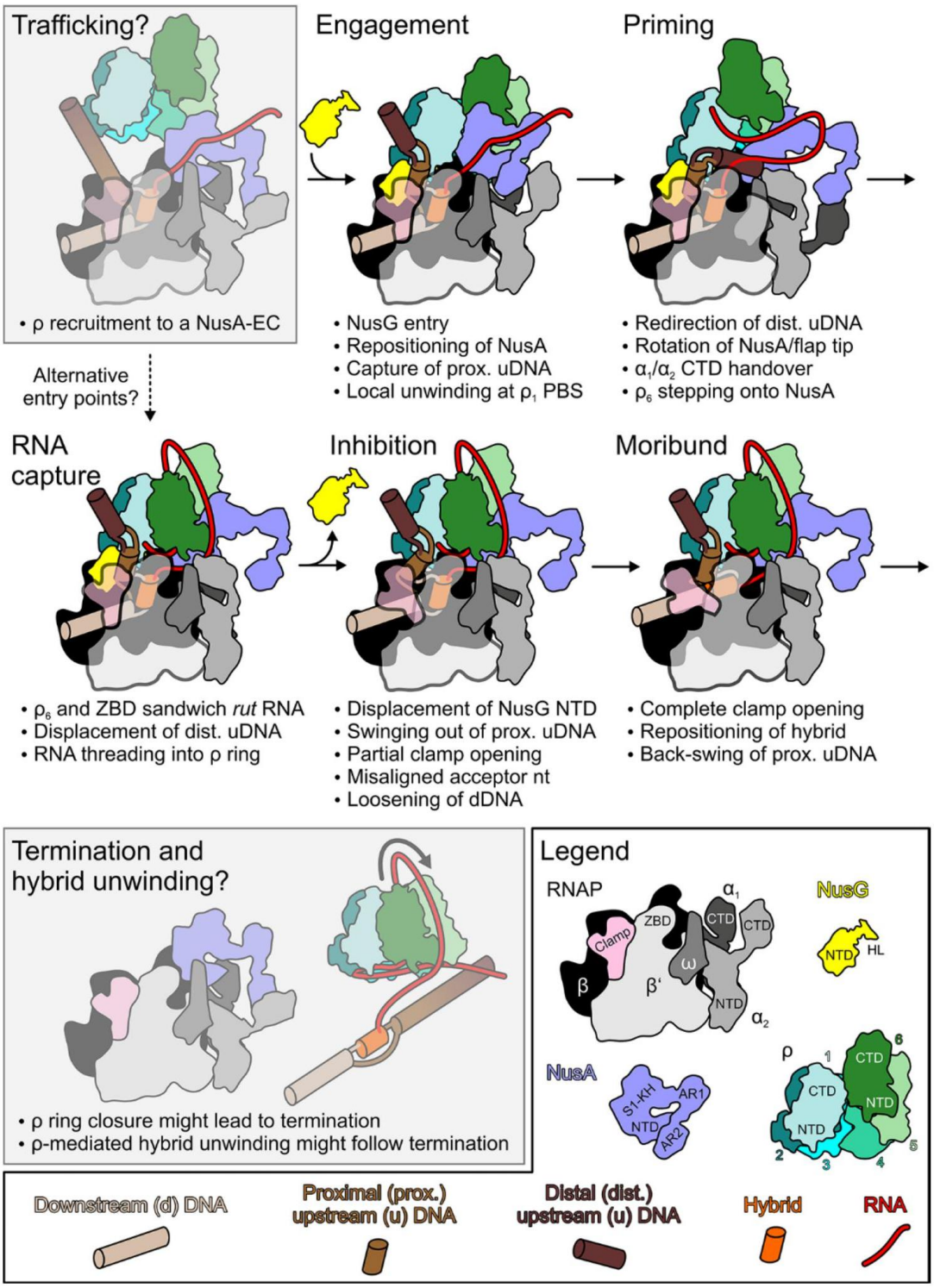


Figure 1.10 Model for an EC-dependent ρ-mediated termination pathway. Trafficking and termination/ hybrid unwinding correspond to hypothetical steps (behind semi-transparent gray boxes) preceding and following the stages resolved by cryoEM in this work. Legend on the lower right and bottom. Coloring as in structural figures except: DNA, upstream to downstream progressively lighter brown; hybrid, orange. (Said, N., Wahl, M.C., et al. 2020)

2.3.4 ATPase HelD regulated RNAP recycling in *Bacillus subtilis*

Bacillus subtilis HelD was initially characterized as an elongation factor, a helicase associated with RNAP complexes (Delumeau, O. et al., 2011). Based on sequence homology, HelD belongs to the DNA and RNA helicases superfamily I. The best characterized helicases, belonging to the same superfamily but only distantly related to HelD, are UvrD and Rep helicases from *E. coli*, or PcrA helicase from *Geobacillus stearothermophilus*. UvrD, which is originally known as DNA helicase II in *E. coli* (Hickson, I.D. et al., 1983), is the founding member of helicases superfamily I and unwinds DNA in a 3′ to 5′ direction (Matson, S.W. and George, J.W. 1987). HelD is strongly expressed during the exponential phase of growth with a further increase in expression in the stationary phase (Nicolas, P. et al., 2012). Although HelD stimulates transcription in an ATP-dependent manner, HelD hydrolyzes ATP independent of RNAP presence (Wiedermannová, J. et al., 2014). HelD and the δ subunit of RNAP work together to promote more efficient recycling of RNAP, which stimulates multiple rounds of *in vitro* transcription (Wiedermannová, J. et al., 2014). The absence of HelD prolongs the lag phase of growth in a similar manner as the absence of δ does; this phenotype negatively affects the ability of the cell to rapidly react to nutritional changes in the environment (Wiedermannová, J. et al., 2014).

2.5 Aims of the study

Transcription has been extensively studied for decades. In the recent years, with the resolution revolution of cryoEM, the structural studies of large bio-macromolecular complexes become much more convenient. Although numerous different functional states RNAPs structures have been unveiled, most of these studies are focused on initiation, elongation, pausing, antitermination and termination, while post-termination has rarely been investigated.

RNAP can get stuck on the DNA template after termination (Kang, W. et al., 2020; Bellecourt, M.J. et al., 2019; Harden, T.T. et al., 2020), RNAP can also form binary complexes with nascent RNA (López de Saro, F.J. et al., 1995; Altmann, C.R. et al., 1994). These stalled RNAPs are unproductive, threatening genome stability and limiting free enzyme pools, or enter hibernating states. The stuck DNA/RNA have to be displaced to recycle RNAP, and ordered recycling is considered an integral phase of the duty cycle of many molecular machines, such as ribosomes (Gohara, D.W. and Yap, M.F. 2018). However, recycling has not garnered similar attention in bacterial transcription to date. Even though several accessory factors could facilitate RNAP detachment from nucleic acids, including σ (Altmann, C.R. et al., 1994; Busby, S. et al., 1981), transcription repair coupling factor Mfd (Selby, C.P. 2017), ρ (Washburn,

R.S. and Gottesman, M.E. 2011) and the NTPase RapA (Sukhodolets, M.V. et al., 2001), they release trapped RNAP under specific circumstances rather than act as genuine recycling factors.

The intrinsic termination is the major form of transcription termination in *B. subtilis,* and RNAP mostly gets stuck on DNA after intrinsic termination. Transcription and translation is uncoupled in *B. subtilis*, the nascent mRNA thus should be displaced from RNAP to participate in translation. Therefore, there should be an efficient way to dismantle DNA/RNA to recycle RNAP in *B. subtilis*. The ATPase HelD, in cooperation with the small subunit δ of RNAP, can stimulate multiple rounds of transcription *in vitro* (Wiedermannová, J. et al., 2014). This effect is probably related to the more efficient recycling of RNAP, since δ is involved in RNAP recycling. However, how exactly the process is performed is unknown. Thus, the overall goal of my thesis is to elucidate the structure of the *B. subtilis* RNAP-δ-HelD complex via biochemical approaches in order to understand a post-termination process: RNAP recycling. Specific aims include:

1. Establishment of an efficient purification protocol of *Bacillus subtilis* RNAP, as we purify the endogenous polymerase, which has a very low yield. Purification of recombinantly produced *Bacillus subtilis* HelD, δ and δ^{NTD} to analyze DNA/RNA displacement in RNAP recycling. And purification of the general transcription initiation factor σ^{A} to analyze HelD release in RNAP recycling.

2. Identification of contents in the purified endogenous RNAP via fingerprinting mass spectrometry.

3. Experimental definition of stable complexes formed between RNAP, HelD, δ and the DNA-RNA hybrid, which mimics the DNA template and the nascent RNA. Establishment of a purification protocol of the stable complex to yield sufficient amounts of highly homogeneous material for structural and functional studies.

4. Determination of the RNAP-δ-HelD complex structure via single particle analysis (SPA) cryoEM.

5. Verification of intra-molecule and inter-molecule contacts in the RNAP-δ-HelD complex and the configuration rearrangement of RNAP that is induced by HelD-δ engagement via cross-linking coupled to mass spectrometry (CLMS).

6. Functional analyses, such as nucleic acid displacement assay, interaction assays and HelD release assays

3 Materials and Methods

3.1 Materials

3.1.1 Chemicals

Standard chemicals, organic substances and solvents, which are not listed here, were ordered by one of the following companies: Sigma-Alderich, Merck, and Roth.

Table 3.1: Chemicals

Chemicals	Company
2-log DNA ladder	New England Biolabs, Germany
1,4-Dithiothreitol	Roth, Germany
2-Mercaptoethanol	Roth, Germany
2-Propanol	Merck, Germany
Acetic acid	Merck, Germany
Acetonitrile	Roth, Germany
Acrylamide solution	Roth, Germany
Ampicilin, sodium salt	Fluka, Switzerland
Calcium chloride dihydrate	Merck, Germany
Chloramphenicol	Roche, Germany
Coomassie brillant blue G-250	Serva, Germany
Coomassie brillant blue R-250	Serva, Germany
Erythromycin	Sigma-Aldrich, Germany
Ethanol	Merck, Germany
Ethylenediamine tetraacetic acid, disodium salt dihydrate	Roth, Germany
Formic Acid	Roth, Germany
Glutathione (reduced)	Sigma-Aldrich, Germany
Glycerol	Merck, Germany
Glycoblue	Thermo Fisher Scientific, USA
D+-Glucose	Aplichem, Germany
Hydrochloric acid	Merck, Germany
Imidazole	Merck, Germany
Isopropanol	Merck, Germany
Kanamycine sulfate	Roth, Germany
Lithium chloride	Fluka, Switzerland
Magnesium acetate	Fluka, Switzerland
Magnesium chloride hexahydrate	Fluka, Switzerland

Magnesiumsulfate	Roth, Germany
N, N, N', N'-Tetramethylethylendiamin (TEMED)	Sigma-Alderich, Germany
n-octylglucoside	Sigma-Alderich, Germany
Potassium acetate	Roth, Germany
Potassium chloride	Merck, Germany
Potassium hydroxide	Merck, Germany
Precision Plus Protein™ Prestained Standards	Bio-rad, Germany
Roti-Phenol/Chloroform	Roth, Germany
Rotiphorese Gel 40 (37.5:1)	Roth, Germany
Rotiphorese Gel 40 (19:1)	Roth, Germany
Sodium acetate	Roth, Germany
Sodium azide	Roth, Germany
Sodium carbonate	Merck, Germany
Sodium chloride	Merck, Germany
Sodium dodecylsulfate (SDS)	Serva, Germany
Sodium hydroxid	Merck, Germany
Sulfo-SDA	Sigma-Alderich, Germany
Trifluoroacetic acid	Roth, Germany
TRIS Acetate	Sigma-Alderich, Germany
TRIS-(hydroxymethyl)-aminomethane (TRIS)	Roth, Germany
Triton X-100	Merck, Germany
Tween -20	Merck, Germany
Urea	Merck, Germany
Tryptone	Roth, Germany
Yeast extract powder	Roth, Germany

3.1.2 Medium components

All mediums were prepared with Milli-Q water and autoclaved.

Table 3.2: Mediums

Media	Composition
50x5052	25 % glycerol, 2.5 % glucose, 10 % α-lactose
50xM	1.25 M Na_2HPO_4, 1.25 M KH_2PO_4, 2.5 M NH_4Cl, 0.25 M Na_2SO_4
1000x Trace Metals	50 mM $FeCl_3$, 20 mM $CaCl_2$, 10 mM $MnCl_2$, 10 mM $ZnSO_4$, 2 mM $CoCl_2$, 2 mM $CuCl_2$, 2 mM $NiCl_2$, 2 mM Na_2MoO_4, 2 mM Na_2SeO_3, 2 mM H_3BO_3 in ~60 mM HCl
Phosphate buffer	0.17 M KH_2PO_4, 0.72 M K_2HPO_4
LB-Medium	1 % Tryptone, 0.5 % *yeast* extract, 0.5 % NaCl
TB Medium	2 % Tryptone, 2.4 % *Yeast* extract, 0.4 % (v/v)

	Glycerol, 10 % (v/v) Phosphate buffer
ZYM-5052	950 ml ZY media supplemented with 2 ml of 1 M $MgSO_4$, 200 µl trace metals (1000 x solution), 20 ml of 50x 5052, 20 ml 50x M

3.1.3 Buffers and solutions

All buffers were prepared with Mili-Q water and filter-sterilized or autoclaved. Adjustments of pH values were performed by the addition of buffered stock solutions or by titration with 37 % HCl and NaOH stock solutions (1 M - 10 M).

Table 3.3: Buffers and stock Solutions

Buffers/Solutions	Composition
5x DNA loading buffer	1x TBE, 30 % (v/v) glycerol, 0.05 % (w/v) xylene cyanol, 0.05 % (w/v) bromphenol blue
1x RNA loading dye	0.5 mg/ml urea, 2 mg/ml xylene cyanol, 2 mg/ml bromophenol blue, 0.1 M EDTA, pH 8.0
1x native RNA loading dye	2 mg/ml xylene cyanol, 2 mg/ml bromophenol blue, 0.1 M EDTA, pH 8.0
10 % Urea gel	19.2 g urea, 10 ml Rotiphorese Gel 40 (19:1), 4 ml 10x TBE in 40 ml H_2O
5x SDS loading buffer	250 mM Tris pH 6.8, 8 %(w/v) SDS, 10 % (v/v) ß-ME, 30 % (v/v) glycerol, 0.02 % (w/v) bromphenol blue
10x TBE	0.89 M Tris pH 8.0, 0.89 M boric acid, 25 mM EDTA
10x SDS running buffer	200 mM Tris pH 6.8, 2 mM glycine, 1 % (w/v) SDS
Stacking gel buffer	0.5 M Tris pH 6.8, 0.4 % (w/v) SDS
Resolving gel buffer	1.5 M Tris pH 8.8, 0.4 % (w/v) SDS
Coomassie staining solution	0.025 % (w/v) Coomassie (R250), 0.025 % (w/v) Coomassie (G250), 30 % (v/v) isopropanol, 7.5 % (v/v) acetic acid
Distaining solution	10 % (v/v) acetic acid

3.1.4 Consumables

Table 3.4: Consumables

Items	Company
Amicon Ultra Concentrators	Millipore, USA
Autoradiography films BioMax MR	Kodak, USA

Dialysis membrane	Spectra/Por, USA
Electroporation cuvettes	Bio-Rad, Germany
Eppendorf safe-lock micro test tubes	Eppendorf, Germany
Falcon tubes (5, 15, 50 ml)	Greiner, Germany
Gloves, Dermatril	KCL, Germany
Needles	Henke Sass Wolf, Germany
Parafilm	Pechiney Plastic Packaging, USA
Quantifoil R1/2 holey carbon grids	Electron Microscopy Sciences, UK
Reaction tubes (0.5; 1.5; 2 ml)	Greiner-Bio-One, Germany
Sterile filters (0.2; 0.45 µm)	Sarstedt, Germany
Syringes	Braun, Germany
Vivaspin concentrators	Vivascience, Sartorius
Whatman 3 MM paper	Whatman plc, UK

3.1.5 Chromatographic resins and columns

Table 3.5: Chromatographic resins and columns

Matrix	Company
Ni-NTA agarose	Macherey-Nagel
Glutathione resin	Macherey-Nagel
MonoQ 5/50 GL, Superdex 200 16/60, Superdex 200 10/30, Superdex 200 PC 3.2/30, Superdex 75 16/60, Superdex 75 10/30, Superose6 PC 3.2/30 Hiprep Heparin FF, Superdex Peptide 3.2/300.	GE Healthcare, USA
EASY-Spray C18 LC column	Thermo Fisher Scientific, USA

3.1.6 Nucleotides

Table 3.6: Nucleotides

Nucleotides	Company
Deoxynucleoside-5'-triphosphate (dNTPs, 100 mM each) (dATP, dCTP, dGTP, dTTP)	New England Biolabs, Germany
Nucleoside-5'-Triphophate (rNTPs, 100 mM) (ATP, CTP, GTP, UTP)	Jena Biosciences, Germany
ATP analogs (ATPγS, AMPPNP, ADP, AMP)	Jena Biosciences, Germany
[α-^{32}P] ATP, [γ-^{32}P] ATP (6000 Ci/mmol, 10 mCi/ml)	Perkin Elmer, Germany

3.1.7 Commercial kits

Table 3.7: Commercial kits

Kits	Company
QIAprep Spin Miniprep Kit	Qiagen, Germany
QIAquick Gel/PCR Purification Kit	Qiagen, Germany

3.1.8 Instrumentation

Table 3.8: Instrumentation

Device	Company
Äkta Prime, Purifier, Micro	Amersham, Freiburg
Allegra X-15R	Beckman Coulter, Germany
Autoclaves	H+P Labortechnik, Germany
Avanti J-26 XP	Beckman Coulter, Germany
DAWN TREOS multi-angle light scattering	Wyatt Technology Corporation, USA
Electrophoresis chamber	Bio-Rad, Germany
Eppendorf Centifuge 5415R	Eppendorf, Germany
Eppendorf Centifuge 5427R	Eppendorf, Germany
FEI Vitrobot Mark IV	Thermo Fisher Scientific, USA
FEI Titan Krios G3i (300 kV) with a Falcon 3EC camera	Thermo Fisher Scientific, USA
Gel documentation system	Bio-Rad, Germany
Geldryer model 583	Bio-Rad, Germany
Gel electrophoresis equipment	Bio-Rad, Germany
Glass-ware	VWR International, Germany
Heating blocks	Eppendorf, Germany
Head over tail Rotor 7637-01	Cole-Parmer, USA
HPLC system	Agilent Technologies, USA
Ice machine	Ziegra, UK
Incubator BK-600	Heraeus, Hanau
Incubation shaker Multitron	Infors, Switzerland
Milli-Q synthesis A10	Milipore, USA
Magnetic stirrer	IKA, Germany
Microliter syringes	Hamilton, Switzerland
Microwave oven	Severin, Germany
NanoDrop 2000 Spectrometer	Thermo Scientific, USA
Orbitrap Fusion Lumos Tribrid mass spectrometer	Thermo Scientific, USA
PCR cycler Peqstar 2x gradient	PeqLab, Germany

pH Meter, Professional Meter PP-20	Sartorius, Germany
Phosphorimager Typhoon 8600	GE Healthcare
Photometer DU 530	Beckmann, Germany
Pipettes, Pipetman	Eppendorf, Germany
Q Exactive HF mass spectrometer	Thermo Scientific, USA
Sonoplus ultrasonic homogenizer HD3100	Bandelin, Germany
Power supplies	Bio-Rad, Germany
Power supply-EPS 3500	Pharmacia Electrophoresis, Germany
Speed Vac Concentrator 5301`	Eppendorf, Germany
Table centrifuge 5415R	Eppendorf, Germany
Thermomixer comfort	Eppendorf, Germany
RefractoMax 520 refractive index detectors	Wyatt Technology Corporation, USA
Tunair flasks	Signa-Alderich, Germany
Ultracentrifuge	Sorvall/Beckman, USA
Ultimate 3000 RSLCnano system	Dionex, Thermo Fisher Scientific, USA
UVP CL‐1000 UV Crosslinker	UVP Inc. USA
VWR Micro STAR17 centrifuge	VWR International, Germany
X-ray film developer X-Omat 2000	Kodak, USA

3.1.9 Enzymes and proteins

Table 3.9: Enzymes and proteins

Enzymes and proteins	Company
Cloned Pfu DNA Polymerase (10 U/µl)	Agilent Technologies, Germany
DNase I (RNase free; 20 U/ul)	Thermo Scientific, Germany
Pyrophosphatase (PPase), inorganic (200 U/mg)	Sigma-Alderich, Germany
Q5 High-Fidelity DNA Polymerase (5000 U/µl)	New England Biolabs, Germany
Restriction endonucleases	New England Biolabs, Germany/ Thermo Scientific, Germany
PreScission protease	Home-made, recombinat
TEV protease	Home-made, recombinat
T4 Polynucleotide Kinase	New England Biolabs, Germany
T4 DNA ligase	New England Biolabs, Germany
T4 DNA polymerase	New England Biolabs, Germany

3.1.10 RNA and DNA oligonucleotides

Synthetic RNA was purchased from IBA-lifesciences, Germany. All synthetic DNA were purchased from Eurofin MWG Operon, Ebersberg, Germany. The oligonucleotides were dissolved in Milli-Q water to a stock concentration of 100 pmol/µl if not otherwise stated and stored at - 20 °C.

Table 3.10: RNA and DNA Oligonucleotides

PCR primers	
helD	
Forward	5'-TGCGCGGATCCATGAATCAGCAGGATAAGGAATGGAAGG-3'
Reverse	5'-TACCGCTCGAGTCATTCAGCAATCTGATATAAGTGAGG-3'
rpoE	
Forward	5'-GTCGGCCATGGGTATCAAACAATATTCACAGG-3'
Reverse	5'-GAGCCAAGCTTATTTAATTTCCTCTTCTTCATCATC-3'
rpoE1-90	
Forward	5'-CCCATGGGGATGGGTATCAAACAATATTCACAG-3'
Reverse	5'-CCCTCGAGTTCTTCATCAAGCTGATCATAAG-3'
SigA	
Forward	5'-CCCATGGGGATGGCTGATAAACAAACCC-3'
Reverse	5'-CCAAGCTTTTATTCAAGGAAATCTTTCAAACG-3'
Scaffold	
ntDNA	5'-GCCGAGCAGCGTAGCATTACTTGTGAGCGGATAAC-3'
tDNA	5'-GTTATCCGCTCACAATGCCACACGCGCTGCTCGGC-3'
RNA	5'-CGUGUGGCA-3'

3.1.11 Plasmids

Table 3.11: Plasmids

Plasmids	**Description**	**Source**
pETM11	Expression vector with a TEV protease-cleavable N-terminal His_6 tag in *E. coli*; KanR	EMBL, Germany
pGEX-6P-1	Expression vector with a PreScission protease cleavable N-terminal GST tag in *E. coli*; AmpR	GE Healthcare, Germany
pGEX-6P-1-bsHelD	N-GST, PreScission site, Gene cloned via *BamHI* and *XhoI*	This study
pETM-11-bsDelta	N-His_6, TEV site, Gene cloned via *NcoI* and *HindIII*	This study
pETM-11-bsDelta$^{1-90}$	N-His_6, TEV site, Gene cloned via *NcoI* and *XhoI*	This study
pETM-11-bsSigA	N-His_6, TEV site, Gene cloned via *NcoI* and *HindIII*	This study

3.1.12 Bacterial strains

Table 3.12: Bacterial strains

Bacterial strains	Description	Source
B. subtilis MH5636	β' subunit with a C-terminal His_{10}-tag,	Qi, Y. and Hulett, F. M. 1998
B. subtilis LK782	β' subunit with a C-terminal His_{10}-tag, *helD::MLS,*	Wiedermannová, J., Krásny, L. et al. 2014
B. subtilis LK1032	β' subunit with a C-terminal His_{10}-tag, *helD::MLS, rpoE::Kan*	Wiedermannová, J., Krásny, L. et al. 2014
E. coli Rosetta (DE3)	Expression for GST-HelD	Novagen, USA
E. coli DH5α	Routine cloning applications	Invitrogen, USA
E. coli BL21(DE3)-RIL	Expression for His_6-δ, His_6-$δ^{NTD}$ and His_6-$σ^A$,	Novagen, USA

3.1.13 Software

Table 3.13: Software

Software	Reference
Coot	Emsley 2010
Corel Draw	Coral Corporation, USA
Image-Quant	GE Healthcare
xiSEARCH	Rappsilber lab, Germany
ProteoWizard	Chambers, M.C. et al. 2012
Pymol	Schrödinger LLC, USA
Phenix	D.Liebschner, et al. 2019
Windows Microsoft office	Microsoft, USA

3.2 Cloning and protein purification

Commercial Kits were used according to the manufacturer's protocol.

3.2.1 Construction of HelD, $σ^A$, δ and $δ^{NTD}$

B. subtilis genomic DNA was extracted from *B. subtilis* strain MH5636. A DNA fragment encoding *B. subtilis* HelD was PCR-amplified from strain MH5636. The PCR product was inserted into expression vector pGEX-6p-1 via *BamHI* and *XhoI* restriction sites, in frame with a region encoding an N-terminal GST-tag. DNA fragments encoding *B. subtilis* $σ^A$, δ or $δ^{NTD}$ were PCR-amplified from strain MH5636 and inserted into pETM-11 vector via *NcoI/HindIII* or *NcoI/XhoI* restriction sites,

respectively, in frame with a region encoding an N-terminal His_6-tag. We confirmed these constructs with DNA sequencing, all of them were successfully cloned to *E. coli* expression vectors.

3.2.1.1 Standard PCR amplification

To clone *B. subtilis* HelD, δ, δ^{NTD} and σ^A, polymerase chain reaction (PCR) was used for amplification of genes or gene fragments from *B. subtilis* genomic DNA with primers listed in Table 3.10. Primers were designed to introduce compatible restriction enzyme cleavage sites at both ends and usually contained 3 additional bases at their 5' ends to allow for efficient digestion by restriction enzymes. In standard cloning PCR reactions, the Q5 High-Fidelity DNA polymerase (NEB) was used according to the manufactures protocol. PCR reactions and cycling programs are shown below:

Parameter	Condition
5x Q5 buffer	1x
dNTPs mixture (10 mM each)	0.2 mM
DNA template (plasmid DNA)	10 - 50 ng
Oligonucleotides (10 µM)	0.5 µM each
Q5 DNA Polymerase	0.1 – 0.2 U/µl
Denaturing temperature	98 °C
Annealing temperature	Tm – 3 °C
Extension time	20 - 30 sec per kb
Extension temperature	72 °C
Number of cycles	35

3.2.1.2 Agarose gel electrophoresis, gel extraction and PCR purification

Analysis of PCR products and gel extraction was performed using agarose gel electrophoresis. PCR products were mixed with 5x DNA loading buffer and loaded to 1 % agarose gel. Commercial DNA ladders were used as a size marker and the agarose solution was supplemented with 0.5 µg/ml ethidium bromide (EtBr) to be able to visualized by UV. The agarose gel was run in 1x TAE buffer at 150 V for 40 min and visualized by UV illumination at 254 nm. The nucleic acid band of the desired size was excised from the gel and extracted using QIAquick gel extraction kit according to the manufacturer's protocol. For PCR purification, the QIAquick PCR purification kit was used according to the manufacturer's protocol.

3.2.1.3 Enzyme digestion and ligation

PCR products and plasmids, harboring the desired restriction site were digested with appropriate restriction enzymes (*BamHI/XhoI* for HelD, *NcoI/HindIII* for σ^A and δ, *NcoI/XhoI* for δ^{NTD}), which produced compatible ends for the subsequent ligation reaction. Reactions were performed according to the manufacturer's protocol. For standard digestion reactions, 1 – 5 µg DNA and 1 U restriction enzyme per 1 µg DNA were used. Digestion products were either purified by agarose gel electrophoresis or PCR products purification. For ligation, typically 100 ng digested plasmid were mixed with two to seven-fold molar excess of the insert DNA and 800 units of T4 DNA ligase in a 20 µl reaction. Reactions were incubated for 1 h at 25 °C or overnight at 4 °C.

3.2.1.4 Transformation of *Escherichia coli* cells

Electro or chemically *E. coli* competent cells used for cloning and protein production were prepared by lab technicians. For transformation *E. coli* competent cells were thawed on ice. For electroporation, the ligation reactions were dialyzed against ddH_2O to remove solute in the reaction. 50 - 100 ng of plasmid or ligation reaction were mixed with 50 - 100 µl electro-competent *E. coli* cells on ice. The mixture was transferred to an ice-cold electroporation cuvette and subjected to a 4.8 ms pulse at either 1.8 kV (1 mm electrode gap) or 2.4 kV (2 mm electrode gap). For chemically *E. coli* competent cells transformation, 50 - 100 ng of plasmid or ligation reaction were mixed with 50 - 100 µl chemically competent *E. coli* cells on ice for 30 min. The mixture was subjected to heat shock in a 42 °C water bath for 45 s, afterwards the mixture was immediately transferred on ice for 2 min. Cells were rescued by addition of 400 µl LB solution and incubated for 1 h at 37 °C in a shaker. Subsequently, the cells were harvested by centrifugation at 1500 x g, resuspended in 100 µl LB medium and streaked out on an LB-agar plate containing the selective antibiotics. Plates were incubated at 37 °C overnight.

3.2.1.5 Plasmids isolation from *Escherichia coli* cells

A single colony from an LB-agar plate, grown overnight at 37 °C, was used to inoculate 1 x LB medium for large-scale purification of plasmids. Plasmid purification was carried out using commercial (Miniprep) kits, according to the manufacturer's protocol.

3.2.1.6 DNA sequencing

All constructed plasmids were tested for containing the desired insert by cloning PCR.

Clones, which showed the expected size product in agarose gel analysis, were verified by DNA sequencing (Seqlab, Göttingen).

3.2.1.7 Determination of nucleic acid concentration

To determine the concentration of nucleic acids, the absorption of an aqueous solution was measured at a wavelength of 260 nm by using a Nanodrop spectrophotometer. The concentration was subsequently calculated by using theoretical values at 260 nm: 1 OD^{260} = 50 µg/ml (ds DNA), 1OD^{260} = 33 µg/ml (ss DNA), and 1 OD^{260} = 40 µg/ml (ss RNA).

3.2.1.8 RNA polyacrylamide gel electrophoresis (urea gels)

For RNA analysis, fractions from binding assays were separated on a denaturing 8 M urea 10% polyacrylamide gel (Rotiphorese 40 (19:1)). Samples were mixed with 1x RNA loading dye, incubated for 5 min at 96 °C and subsequently loaded to the gel. Gels were run at 200 V in 1x TBE, stained in 1x TBE solution supplemented with 1 µg/ml EtBr and visualized by UV illumination at 254 nm. For the native RNA PAGE, urea was not used.

3.2.2 Protein production

3.2.2.1 Antibiotics for *Escherichia coli* and *Bacillus subtilis* cultivation

The following concentrations of antibiotics were used in standard cultivation conditions of *E. coli* or *B. subtilis* cells, if not otherwise stated:

- ampicillin 100 µg/ml
- chloramphenicol 34 µg/ml
- kanamycin 50 µg/ml
- erythromycin 15 µg/ml

3.2.2.2 Protein expression in *Escherichia coli*

B. subtilis HelD, σ^A, δ and δ^{1-90} expressions were performed in ZYM-5052 (50 mM phosphate) auto-induction complex medium (Studier, F.W. 2005) with the suitable

antibiotics. Auto-induction is based on a metabolic regulation mechanism that allows bacterial cells to selectively use a higher energetic carbon sources (i.e. glucose) over an alternative, lower energetic carbon sources (i.e. lactose) in the presence of both. To ensure bacterial growth to a saturated level before protein induction, lactose - which induces expression of T7 polymerase that in turn transcribes the plasmid coded open-reading-frame (ORF) of interest - is prevented from bacterial uptake by addition of higher energetic carbon sources, such as glucose, glycerol and amino acids. After depletion of these nutrient sources, lactose is taken-up and converted into allolactose, the natural inducer of the lac operon. Subsequently, the T7 lac promoter regulated T7 RNA polymerase gene is induced, which causes T7 RNA polymerase-dependent expression of plasmid-coded ORFs. Preparative protein production was performed in Tunair flasks, which maintained adequate aeration during cell growth.

For protein production, plasmids containing B.s HelD gene were transformed into *E. coli* Rosetta2(DE3) cells, and plasmids containing B.s σ^A, δ and $\delta^{1\text{-}90}$ were transformed into *E. coli* BL21(DE3)-RIL cells. The cells were grown in auto-induction medium to an OD^{600} of 1.0 at 37 °C, and then incubated for an additional 16 h at 20 °C, to an approximate OD^{600} of 11.

Optical densities of cultures were measured in plastic cuvettes in an Ultrospec 3000 pro spectral photometer at 600 nm wavelength.

3.2.2.3 *Bacillus subtilis* RNAP production

B. subtilis strains MH5636, LK782 (*ΔhelD*) or LK1032 (*ΔhelDΔrpoE*; Table 3.12) were used to produce stationary phase RNAP, $RNAP^{\Delta HelD}$ or $RNAP^{\Delta\delta\Delta HelD}$, respectively. Chromosomes in these strains are engineered to produce a β' subunit with a C-terminal His_{10}-tag. Strains were grown in TB medium with correlated antibiotics at 37 °C to an OD^{600} of 1.0 and were then shifted to 18 °C and grown to an OD^{600} of about 11.

3.2.3 Protein purification

For all protein preparations, cells were harvested at 4 °C and resuspended in the respective binding buffer. Cells were lysed by sonication at 70 % energy for 1 sec interrupted by 2 sec pause for a total of 20 min and cell debris were removed by centrifugation.

3.2.3.1 Purification of *Bacillus subtilis* RNAP

B. subtilis cells were harvested by centrifugation, resuspended in buffer A (50 mM Na_2HPO_4, 300 mM NaCl, 3 mM 2-mercaptoethanol, 5 % [v/v] glycerol, pH 7.9) and

lysed by sonication. The lysate was cleared by centrifugation. RNAP variants were captured on Ni^{2+}-NTA affinity resin, washed with buffer A supplemented with 25 mM imidazole and eluted with buffer A supplemented with 250 mM imidazole. The eluate was dialyzed overnight against 50 mM Na_2HPO_4, 300 mM NaCl, 3 mM DTT, 5 % [v/v] glycerol, pH 7.9, loaded on a 5 ml HiTrap Heparin HP column, washed with buffer B (50 mM TRIS-HCl, 100 mM NaCl, 3 mM DTT, 0.1 mM EDTA, 5 % [v/v] glycerol, pH 7.9) and eluted with a linear gradient to buffer B with 700 mM NaCl. Fractions containing RNAP were pooled and further purified by SEC on a HiLoad Superdex 200 Increase 16/600 column in 20 mM TRIS-HCl, 150 mM NaCl, 0.5 mM DTT, 5 % (v/v) glycerol, pH 8.0. The final samples were concentrated to approximately 16 mg/ml. RNAP produced from strain MH5636 was directly used for EM sample preparation. Other RNAP preparations were aliquoted, flash frozen in liquid N_2 and stored at -80 °C

3.2.3.2 Purification of *Bacillus subtilis* HelD

GST-HelD cells were harvested by centrifugation, resuspended in buffer C (50 mM TRIS-HCl, 500 mM NaCl, 1 mM 2-mercaptoethanol, 10 % [v/v] glycerol, pH 7.9) and lysed by sonication. The lysate was cleared by centrifugation, GST-HelD was captured on glutathione resin, washed with buffer C and eluted with 50 mM TRIS-HCl, 300 mM NaCl, 1 mM DTT, 10 % (v/v) glycerol, 20 mM reduced glutathione, pH 7.9. Eluted fractions were dialyzed against buffer D (20 mM TRIS-HCl, 200 mM NaCl, 1 mM DTT, 5 % [v/v] glycerol, pH 7.9) in the presence of GST-tagged PreScission protease. HelD was separated from uncleaved protein, GST and GST-PreScission by a second passage through glutathione resin. The flowthrough was further purified by SEC on a HiLoad Superdex 200 Increase 16/600 column equilibrated in buffer D. Fractions containing HelD were concentrated, aliquoted, flash frozen in liquid N_2 and stored at -80 °C.

3.2.3.3 Purification of *Bacillus subtilis* δ and $δ^{1-90}$

His_6-δ or His_6-$δ^{NTD}$ cells were harvested by centrifugation, resuspended in 50 mM TRIS-HCl, 500 mM NaCl, 0.5 mM 2-mercaptoethanol 5 % [v/v] glycerol, pH 6.0, and lysed by sonication. The lysate was cleared by centrifugation, His_6-δ/His_6-$δ^{NTD}$ was captured on Ni^{2+}-NTA resin, washed with 50 mM TRIS-HCl, 300 mM NaCl, 0.5 mM 2-mercaptoethanol, 10 mM imidazole, 5 % (v/v) glycerol, pH 6.0, and eluted with 20 mM TRIS-HCl, 150 mM NaCl, 0.5 mM 2-mercaptoethanol, 400 mM imidazole, 5 % (v/v) glycerol, pH 6.0. For the assembly of complexes for cryoEM analysis, eluted His_6-δ was supplemented with His-tagged TEV protease (1:40 [w/w]), dialyzed against buffer E (20 mM TRIS-HCl, 150 mM NaCl, 1 mM DTT, 5 % (v/v) glycerol, pH 6.0) overnight and passed through fresh Ni^{2+}-NTA resin to remove uncleaved His_6- δ, cleaved His_6-tag His-tagged TEV protease. Proteins were further purified by SEC on a

Superdex75 Increase 10/300 column in buffer E. Fractions containing His_6-δ, δ or His_6-$δ^{NTD}$ were concentrated, aliquoted, flash frozen in liquid N_2 and stored at - 80 °C.

3.2.3.4 Purification of *Bacillus subtilis* $σ^A$

$σ^A$ cells were harvested by centrifugation, resuspended in buffer F (20 mM TRIS-HCl, 500 mM NaCl, 1 mM 2-mercaptoethanol, 5 % [v/v] glycerol, pH 7.5) supplemented with 20 mM imidazole, and lysed by sonication. The lysate was cleared by centrifugation, His_6-$σ^A$ was captured on Ni^{2+}-NTA resin, washed with buffer F supplemented with 50 mM imidazole, and eluted with buffer F supplemented with 400 mM imidazole. Eluted His_6-$σ^A$ was supplemented with His-tagged TEV protease (1:40 [w/w]), dialyzed against buffer F supplemented with 1 mM EDTA overnight and passed through fresh Ni^{2+}-NTA resin to remove uncleaved His_6-$σ^A$, cleaved His_6-tag and His-tagged TEV protease. The target protein was further purified by SEC on a Superdex75 Increase 16/600 column in 25 mM TRIS-HCl, 300 mM NaCl, 0.1 mM DTT, 5 % (v/v) glycerol, pH 7.5. Fractions containing $σ^A$ were concentrated, aliquoted, flash frozen in liquid N_2 and stored at - 80 °C.

3.2.3.5 Determination of the protein concentration

The protein concentration is proportional to the absorbance at a wavelength of 280 nm, where the aromatic amino acids, and in particular tryptophan, absorb. The measured absorbance (A), the theoretical extinction coefficient (ε), and the path-length of the cuvette (d) can be used to calculate the final protein concentration according to the Lambert-Beer law:

$$A = εdc$$

Protein extinction coefficient factors were calculated by submitting the protein amino acid sequence to the ExPASy ProtParam tool (https://web.expasy.org/protparam/).

3.2.3.6 Sodium dodecylsulfate polyacrylamide gel electrophoresis

Denaturing SDS polyacrylamide gel electrophoresis (SDS-PAGE) was used to analyze the molecular weight of the proteins (Laemmli, U.K. 1970). Acrylamide gels of 10 – 15 % (37.5:1 acrylamide: bis-acrylamide) were used according to the size of the proteins to be separated. The gel consists of two parts, a stacking gel (1x stacking buffer, 16 % (v/v) 37.5:1 acrylamide: bis-acrylamide solution) and a separating gel (1x resolving gel buffer and 25 – 40 % (v/v) 37.5:1 acrylamide: bis-acrylamide solution). Gels were polymerized with 0.3 % (w/v) APS and 0.03 % (v/v) TEMED.

Protein samples were denatured in protein loading buffer and incubated at 95 °C for 5 min. Gels were typically run at 250V and stopped when the bromphenol blue exits the bottom of the gel.

3.2.3.7 Gel staining and destaining

After SDS-PAGE, the proteins were visualized by incubating the gels in Coomassie staining solution (Sambrook, J. and Fritsch, E.F. 1989) and distained by washing in 10 % acetic acid H_2O solution.

3.3 Biochemical and structural biology methods

3.3.1 Fingerprinting mass spectrometry sample preparation

The protein SDS gel was placed on a clean surface and the desired bands were cut. The cut bands were further chopped to pieces of approximately 1 mm × 1 mm and transferred to 0.5 ml Eppendorf tubes. The chopped gel pieces were incubated in a 1: 1 mix of acetonitrile and 100 mM NH_4HCO_3. The mixtures were shaken 15 min at room-temperature. Then the supernatant was discarded and replaced by 100 % acetonitrile. The gel pieces were incubated until turning to a whitish-opaque colour. Afterwards, the acetonitrile was removed and the gel pieces were lyophilized. The gel pieces were then incubated with 100 mM DTT in 100 mM NH_4HCO_3 to reduce the disulfide bonds. The mixtures were incubated at 56 °C with shaking. The samples were washed with 100 mM NH_4HCO_3 several times. The gel pieces were then incubated with 100% acetonitrile for 5 - 10 min at room temperature. The samples were then centrifuged and the supernatant was discarded. The gel pieces were then lyophilized. Trypsin solution (1 mg/ml, sequencing grade, Sigma) was added to the samples to a final concentration of 12.5 ng/ul. The reactions were kept on ice for 30 min, and then incubated over-night at 37 °C. The samples were centrifuged and the supernatant containing the digested peptides was collected, termed S1. After having removed S1 supernatant, 10 µl of 40% acetonitrile / 0.1% trifouric acid was added to the gel pieces and incubated at room-temperature for several hours, overnight or even up to 2 days. Afterwards, the supernatant was collected, termed S2. S1 and S2 were both sent for MS analysis.

3.3.2 Cross-linking mass spectrometry

Sulfo-SDA predominantly establishes lysine-X crosslinks through a primary

amine-reactive moiety on one side and a UV-activatable moiety on the other (theoretical crosslinking limit 25 Å). Sulfo-SDA was prepared at 3 mg/ml in 20 mM HEPES-NaOH, 5 mM $Mg(OAc)_2$, 300 mM NaCl, 5 mM DTT, 5% (v/v) glycerol, pH 8.0 immediately prior to addition of $RNAP^{\Delta\delta\Delta HelD}$, $RNAP^{\Delta\delta\Delta HelD}$-δ, $RNAP^{\Delta\delta\Delta HelD}$-HelD or $RNAP^{\Delta\delta\Delta HelD}$-δ-HelD (protein: sulfo-SDA 1:3 [w/w]). Samples were incubated on ice for two hours and then irradiated in a thin film using 365 nm UV irradiation (UVP CL‐1000 UV Crosslinker, UVP Inc.) for 20 min on ice (5 cm distance from UV-A lamp). The crosslinked samples were separated by 4 - 12 % BIS-TRIS NuPAGE, and gel bands corresponding to crosslinked monomeric complexes were excised and digested in-gel as described by Chen, Z.A. et al. 2010. Resulting peptides were desalted using C18 StageTips (Rappsilber, J. et al., 2007).

10 % of each sample were analyzed by LC-MS/MS without fractionation, the remaining 90 % were fractionated using SEC on a Superdex Peptide 3.2/300 column (GE Healthcare) in 30 % (v/v) acetonitrile, 0.1 % (v/v) trifluoroacetic acid at a flow rate of 10 µl/min to enrich for crosslinked peptides (Leitner, A., Walzthoeni, T. and Aebersold, R. 2014). The first six peptide-containing fractions (50 µl each) were collected, solvent was removed using a vacuum concentrator and the fractions were analyzed by LC-MS/MS on an Orbitrap Fusion Lumos Tribrid mass spectrometer (Thermo Fisher Scientific), connected to an Ultimate 3000 RSLCnano system (Dionex, Thermo Fisher Scientific).

The non-fractionated samples were injected onto a 50 cm EASY-Spray C18 LC column (Thermo Fisher Scientific) operated at 50 °C. Peptides were separated using a linear gradient going from 2 % mobile phase B (80 % [v/v] acetonitrile, 0.1 % [v/v] formic acid) to 40 % mobile phase B in mobile phase A (0.1 % [v/v] formic acid) at a flow rate of 0.3 µl/min over 110 minutes, followed by a linear increase from 40 % to 95 % mobile phase B in 11 minutes. Eluted peptides were ionized by an EASY-Spray source (Thermo Fisher Scientific) and MS data were acquired in the data-dependent mode with the top-speed option. For each three-second acquisition cycle, the full scan mass spectrum was recorded in the Orbitrap with a resolution of 120,000. The ions with a charge state from 3^+ to 7^+ were isolated and fragmented using higher-energy collisional dissociation (HCD) with 30 % collision energy. The fragmentation spectra were then recorded in the Orbitrap with a resolution of 50,000. Dynamic exclusion was enabled with single repeat count and 60 s exclusion duration.

SEC fractions were analyzed using an identical LC-MS/MS setup. Peptides were separated by applying a gradient ranging from 2 % to 45 % mobile phase B (optimized for each fraction) over 90 min, followed by ramping up mobile phase B to 55 % and 95 % within 2.5 min each. For each three-second data-dependent MS acquisition cycle, the full scan mass spectrum was recorded in the Orbitrap with a resolution of 120,000. The ions with a charge state from 3^+ to 7^+ were isolated and fragmented using HCD. For each isolated precursor, one of three collision energy settings (26 %, 28 % or 30 %) was selected for fragmentation using a data-dependent decision tree based on the m/z and charge of the precursor. The fragmentation spectra were

recorded in the Orbitrap with a resolution of 50,000. Dynamic exclusion was enabled with single repeat count and 60 s exclusion duration.

LC-MS/MS data generated from the four complexes were processed separately. MS2 peak lists were generated from the raw MS data files using the MSConvert module in ProteoWizard (version 3.0.11729). The default parameters were applied, except that Top MS/MS Peaks per 100 Da was set to 20 and the denoising function was enabled. Precursor and fragment m/z values were recalibrated. Identification of crosslinked peptides was carried out using xiSEARCH software (Mendes, M.L., Rappsilber, J. et al. 2019) (https://www.rappsilberlab.org/software/xisearch; version 1.7.4). For $RNAP^{\Delta\delta\Delta HelD}$, peak lists were searched against the sequence and the reversed sequence of RNAP subunits (α, β, β' and ε) and two co-purified proteins, σ^A and σ^B. For $RNAP^{\Delta\delta\Delta HelD}$-δ, $RNAP^{\Delta\delta\Delta HelD}$-HelD and $RNAP^{\Delta\delta\Delta HelD}$-δ-HelD samples, protein sequences of δ, HelD or both were included in the database. The following parameters were applied for the search: MS accuracy = 4 ppm; MS accuracy = 10 ppm; enzyme = trypsin (with full tryptic specificity); allowed number of missed cleavages = 2; missing monoisotopic peak = 2; crosslinker = sulfo-SDA (the reaction specificity for sulfo-SDA was assumed to be for lysine, serine, threonine, tyrosine and protein N termini on the NHS ester end, and any amino acid residue for the diazirine end); fixed modifications = carbamidomethylation on cysteine; variable modifications = oxidation on methionine and sulfo-SDA loop link. Identified crosslinked peptide candidates were filtered using xiFDR (Fischer, L. and Rappsilber, J. 2017). A false discovery rate of 5 % on residue-pair level was applied with the "boost between" option selected. Crosslinked residue pairs identified from the four complexes are summarized in Table 4.3 and Appendix 7.2.

3.3.3 RNAP-HelD-δ complex assembly

Equimolar amounts of tDNA, ntDNA and RNA were mixed in buffer G (20 mM TRIS-HOAc, 5 mM $Mg[OAc]_2$, 100 mM KOAc, 2 mM DTT, 5 % [v/v] glycerol, pH 8.0) and annealed by heating to 95 °C for 5 min and subsequent cooling to 25 °C at 1 °C/min. The annealed scaffold was incubated with *B. subtilis* RNAP in a 1.3:1 molar ratio in buffer H (20 mM TRIS-HOAc, 5 mM $Mg[OAc]_2$, 300 mM KOAc, 2 mM DTT, 5 % [v/v] glycerol, pH 8.0) for 10 min on ice, then for 10 min at 32 °C. Equimolar amounts (to RNAP) of δ and HelD were added stepwise, followed by incubation for 10 min at 32 °C after each addition. The mixture was subjected to SEC on a Superdex 200 Increase 3.2/300 column (GE Healthcare) in buffer H. Fractions containing RNAP, δ and HelD were pooled and concentrated to approximately 5 mg/ml.

3.3.4 CryoEM data collection and processing

Immediately before preparation of the grids, the sample was supplemented with

0.15 % (w/v) n-octylglucoside. 3.8 µl of the final mixture were spotted on plasma-treated Quantifoil R1/2 holey carbon grids at 10 °C/100 % humidity, and plunged into liquid ethane using a FEI Vitrobot Mark IV. Image acquisition was conducted on a FEI Titan Krios G3i (300 kV) with a Falcon 3EC camera at a nominal magnification of 92,000 in counting mode using EPU software (Thermo Fisher Scientific) with a calibrated pixel size of 0.832 Å. A total electron dose of 40 $e^-/Å^2$ was accumulated over an exposure time of 36 s. Movie alignment was done with MotionCor2 (Zheng, S.Q. et al., 2017) using 5x5 patches followed by ctf estimation with Gctf (Zhang, K. 2016).

All following image analysis steps were done with cryoSPARC (Punjani, A. et al., 2017). Class averages of manually selected particles were used to generate an initial template for reference-based particle picking from 9,127 micrographs. Particle images were extracted with a box size of 440 and binned to 110 for initial analysis. *Ab initio* reconstruction using a small subset of particles was conducted to generate an initial 3D reference for 3D heterogeneous refinement. The dataset was iteratively classified into two well-resolved populations representing monomeric and dimeric RNAP-δ-HelD. Selected particles were re-extracted with a box of 220 and again classified in 3D to further clean the dataset. Finally, selected particle images were re-extracted with a box of 280 (1.3 Å/px) and subjected to local refinement using a generously enlarged soft-mask for monomeric or dimeric RNAP-δ-HelD. Local refinement of the dimer particles using the monomeric mask was conducted as a control to trace differences of RNAP-δ-HelD in the authentic monomer and dimer structures. After per-particle CTF correction, non-uniform refinement was applied to generate the final reconstructions.

3.3.5 Model building and refinement

The final cryoEM map for the dimeric RNAP-δ-HelD complex (Figure 4.11 b) was used for initial model building. Coordinates of *Mycobacterium smegmatis* RNAP α, β and β' subunits (PDB ID 5VI8) (Hubin, E.A. et al., 2017) were docked into the cryoEM map using Coot (Emsley, P. et al., 2010). Modeling of δ was based on the NMR structure of *B. subtilis* δ (PDB ID 2M4K) (Papoušková, V. et al., 2013). Modeling of ε was supported by the structure of YkzG from *Geobacillus stearothermophilus* (PDB ID 4NJC) (Keller, A.N. et al, 2014). Model building of HelD was supported by the structure of UvrD helicase from *E. coli* (PDB ID 3LFU) (Jia, H. et al., 2011) as well as the C-terminal domain of a putative DNA helicase from *Lactobacillus plantarun* (PDB ID 3DMN). The subunits were manually rebuilt into the cryoEM map. The model was completed and manually adjusted residue-by-residue, supported by real space refinement in Coot. The manually built model was refined against the cryoEM map using the real space refinement protocol in PHENIX (Afonine, P.V. et al., 2012). Model building of the monomeric complex was done in the same way but starting with a model of half of the dimeric complex. The structures were evaluated with Molprobity

(Williams, C.J. et al., 2018). Structure figures were prepared using PyMOL (Version 1.8 Schrödinger, LLC).

3.3.6 Size exclusion chromatography/multi-angle light scattering

SEC/MALS analysis was performed on an HPLC system (Agilent) coupled to mini DAWN TREOS multi-angle light scattering and RefractoMax 520 refractive index detectors (Wyatt Technology). RNAP-δ-HelD complex was assembled as for cryoEM. 60 µl of the sample at 1 mg/ml were chromatographed on a Superose 6 Increase 10/300 column in buffer H or buffer H plus 0.15 % (w/v) n-octylglucoside, supplemented with 0.02 % (w/v) NaN_3, at 18 °C with a flowrate of 0.6 ml/min. Data were analyzed with the ASTRA 6.1 software (Wyatt Technology) using monomeric bovine serum albumin (Sigma-Aldrich) as a reference.

3.3.7 Negative staining EM analysis

RNAP-δ-HelD complex was prepared as for cryoEM analysis, diluted to 25 µg/ml in buffer H and supplemented with 0.15 % n-octylglucoside or buffer immediately before grid preparation. 4 µl of the samples were added to glow-discharged Formvar/carbon grids (S162, Plano GmbH), left to settle for 40 s and manually blotted with Whatman paper No. 1, followed by addition of 4 µl of 1 % (w/v) uranyl acetate staining solution. After 40 s incubation, the grids were manually blotted and dried at ambient temperature overnight. Samples were imaged on an FEI Talos L120C TEM, operated at 120 kV, equipped with an FEI CETA 16M CCD camera at a nominal magnification of 57,000x. The calibrated pixel size was 1.96 Å/px. Images were acquired manually in low dose mode using TEM Imaging & Analysis (TIA) software, supplied by the manufacturer, accumulating a total electron dose of 50 $e^-/Å^2$. Image analysis was done with cryoSPARC. After CTF estimation, manually selected particle images were used as reference for template-based particle picking. Particle images were extracted with a box size of 160 px and resampled to 80 px. A mask of 220 Å diameter was applied during 2D classification.

3.3.8 Interaction assays

HelD interactions with δ or $δ^{NTD}$ were analyzed by analytical SEC. 21 µM HelD and 42 µM δ or $δ^{NTD}$ were mixed in 20 mM HEPES-NaOH, 50 mM NaCl, 1 mM DTT, pH 7.5, and incubated for 10 min at room temperature. 50 µl of the samples were loaded on a Superdex 200 Increase PC 3.2 column and chromatographed at 4 °C with a flow rate of 40 µl/min. Fractions were analyzed by 12.5 % SDS-PAGE.

3.3.9 Electrophoretic gel mobility shift assays

Equimolar amounts of 5'-[^{32}P]-labeled ntDNA and unlabeled tDNA capable of forming an artificial bubble, or additionally an RNA 9-mer with complementarity to the tDNA in the bubble (Table 3.10), were mixed in buffer G and annealed by heating to 95 °C for 5 min and subsequent cooling to 25 °C at 1 °C/min. 20 nM (to achieve a 10 nM concentration in the final mix) of the labeled DNA duplex or DNA/RNA scaffold were incubated with 2 µM (to achieve a 1 µM concentration in the final mix) RNAP$^{\Delta\delta\Delta HelD}$ in buffer G for 10 min at 4 °C, followed by an additional 10 min incubation at 32 °C. Subsequently, (i) buffer, (ii) HelD (1 µM final concentration); (iii) δ (1 µM final concentration), (iv) combinations of HelD (1 µM final concentration) and δ (titrated final concentration, Figure 4.12) or (v) HelD and δ^{NTD} (1 µM final concentration each) were added, and the samples were further incubated for 10 min at 32 °C. Samples were loaded on a 4 % native PAGE gel and electrophoresed in 0.5X TBE buffer. Radioactive bands were visualized using a Storm phospohorimager and quantified using ImageQuant software.

3.3.10 HelD release assays

Equimolar amounts of HelD and stationary phase RNAP were mixed in buffer I (20 mM TRIS-HCl, 300 mM NaCl, 2 mM DTT, 5 % (v/v) glycerol, pH 8.0), incubated for 10 min on ice and then for 10 min at 32 °C. The sample was chromatographed on a HiLoad Superdex 200 Increase 10/300 column in buffer I. Fractions were analyzed by 12.5 % SDS-PAGE. Fractions containing RNAP-HelD complex were collected and concentrated to approximately 3 mg/ml (6.7 µM). 80 µl 6.7 µM complex were mixed with buffer I, 5 mM Mg^{2+}- ATPγS/AMPPNP/ATP/ADP/AMP, 6.7 µM σ^A or σ^A plus Mg^{2+}-ATPγS in buffer I. 90 µl of the samples were loaded on a Superdex 200 Increase PC 3.2 column and chromatographed at 4 °C with a flow rate of 40 µl/min. Fractions were analyzed by 12.5 % SDS-PAGE.

3.3.11 Structure comparisons

Structures were compared by global superposition of complex structures or by superposition of selected subunits in complexes using the "secondary structure matching" algorithm implemented in Coot or the "align" algorithm implemented in PyMOL.

4 Results

4.1 Purification of *Bacillus subtilis* RNAP

B. subtilis RNAP was purified from *B. subtilis* strain MH5636, expressed with a chromosomally encoded C-terminal His_{10}-tag in the ß' subunit, and purified using a three-step purification protocol (Figure 4.1 a-c).

As shown in Figure 4.1a-c, RNAP can be efficiently purified using endogenously expressed protein. Interestingly, SDS-PAGE analysis after final SEC revealed more complex assembly of the transcription apparatus during stationary phase growth. MS-based fingerprinting analysis of individual bands identified, apart from the core RNAP, sub-stoichiometric amounts of HelD, PriA, the primosomal replication factor Y, and the two general *B. subtilis* sigma factors σ^A and σ^B (Figure 4.1 d).

In order to analyze the structural and functional role of the δ subunit and HelD in context of the RNAP-δ-HelD complex, RNAP variants lacking HelD ($RNAP^{\Delta HelD}$) or δ and HelD ($RNAP^{\Delta\delta\Delta HelD}$) were also purified from *B. subtilis* strains LK782 (*ΔhelD*) and LK1032 (*ΔhelDΔrpoE*), respectively (Figure 4.1 e). The $RNAP^{\Delta HelD}$ shows a marked loss of the ω subunit.

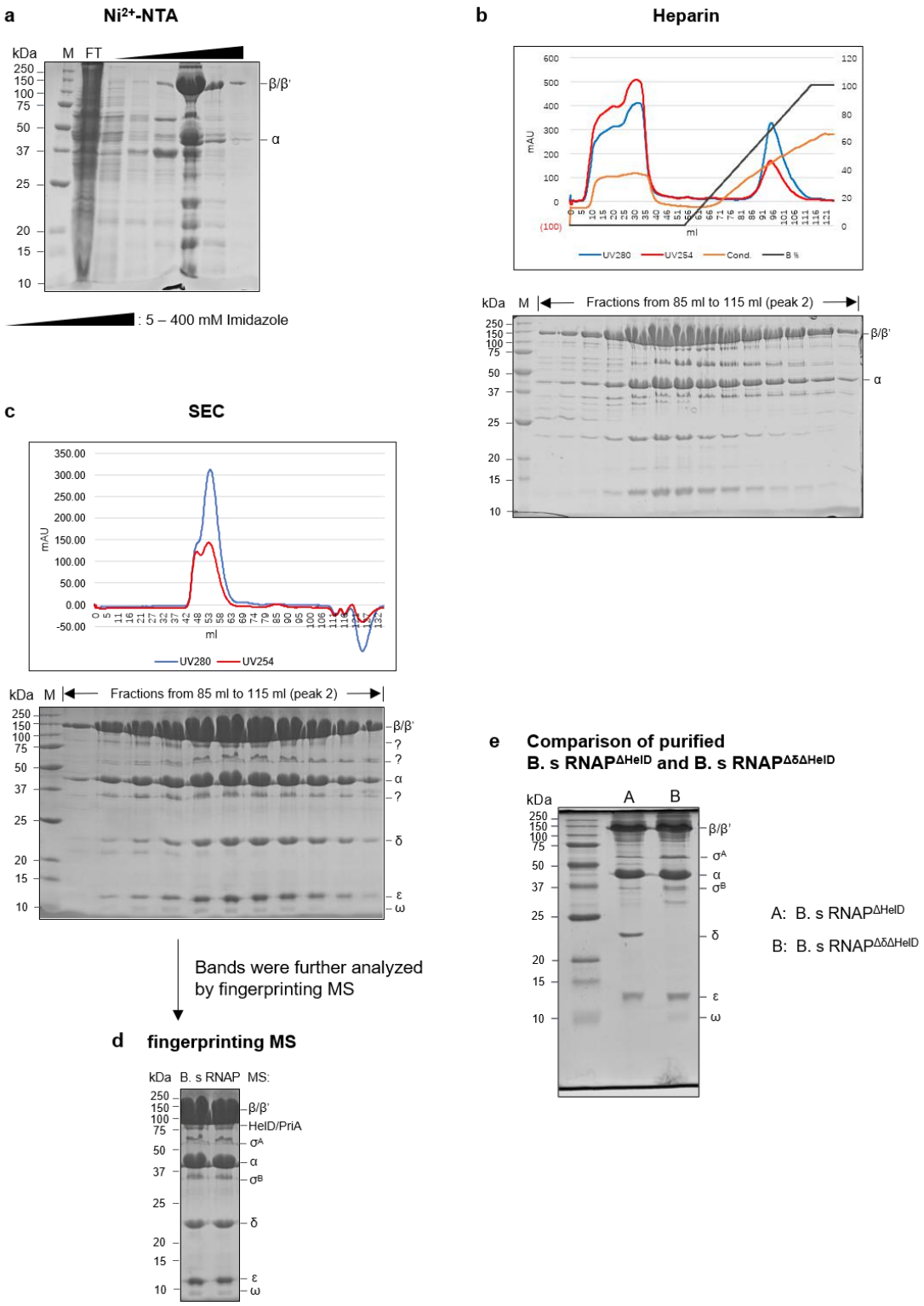


Figure 4.1 a, Affinity purification of Bs RNAP. 12.5% SDS-PAGE analysis of elution fractions after Ni^{2+}-NTA purification. Bands were visualized using 0.025 % (w/v) Coomassie staining solution. b, Ion-exchange chromatography. top: Elution profile of the Heparin column. bottom: Fractions analyzed by 12.5 % SDS-PAGE. c, SEC. top: Chromatogram of the SEC run. bottom: 12.5 % SDS-PAGE analysis of peak fractions from SEC. d, fingerprinting MS, proteins were identified by cutting out the respective bands, in-gel trypsin digestion and MS-based

fingerprinting. e, 13.5 % SDS-PAGE analysis of purified RNAP$^{\Delta HelD}$ and RNAP$^{\Delta\delta\Delta HelD}$

4.2 HelD directly interacts with δ^{CTD}

HelD has been shown to stimulate RNAP recycling with the help of δ (Wiedermannová, J. et al., 2014). δ consist of 173 amino acids and can be roughly subdivided into an N-terminal domain (NTD, residues 1-90) and a C-terminal region (CTR, residues 91-173). To investigate if and how HelD interacts with δ, we recombinantly produced and purified δ, δ^{NTD} and HelD (see “Material and Methods”) and conducted analytical SEC.

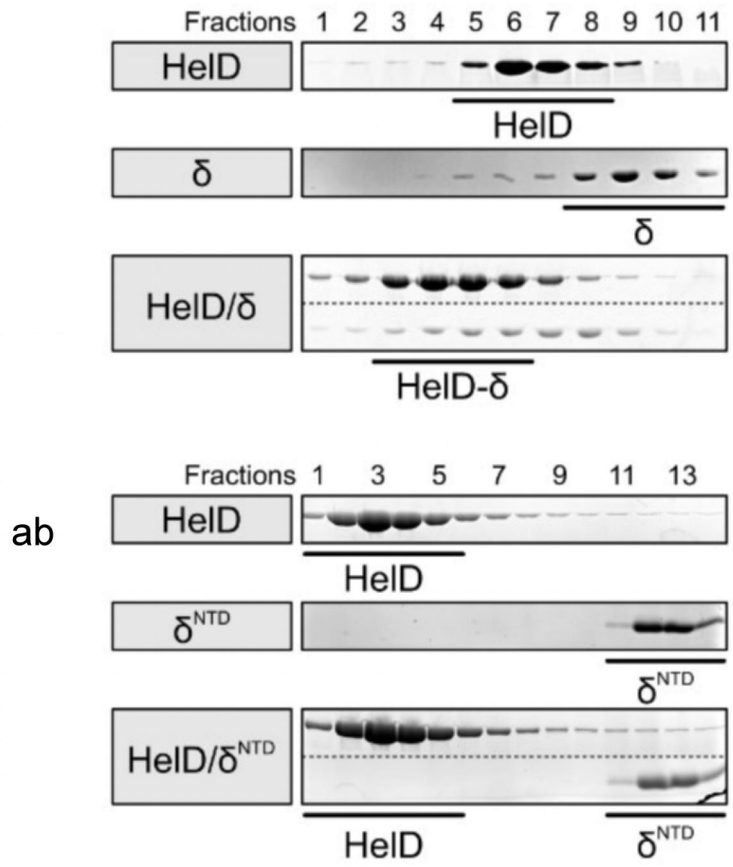


Figure 4.2: HelD interacts directly with δ. Analytical SECs of a HelD/δ mixture (a, lower panel) are compared to a HelD/δ^{NTD} mixture (b, lower panel), and to the isolated proteins (upper two panels). SEC fractions are indicated and analyzed by 12.5 % SDS-PAGE. (with Huang, Y.H.)

In Figure 4.2 a, the upper two lines are SEC runs of individual HelD and δ, respectively, while the lower line is the SEC run of a HelD and δ mixture. As shown in Figure 4.2 a, HelD co-migrated with δ and formed a HelD-δ complex that eluted earlier than the individual proteins. HelD and δ^{NTD}, in contrast, did not show any co-migration effect under the same conditions (Figure 4.2 b). Thus, we could confirm a stable HelD-δ interaction that is directly mediated by the C-terminal region of δ.

4.3 RNAP-δ-HelD complex assembly

To structurally investigate the RNAP-δ-HelD complex, we tested for stable complex assembly between RNAP, δ and HelD. δ and HelD were recombinantly produced and purified as described in the “Materials and Methods” section.

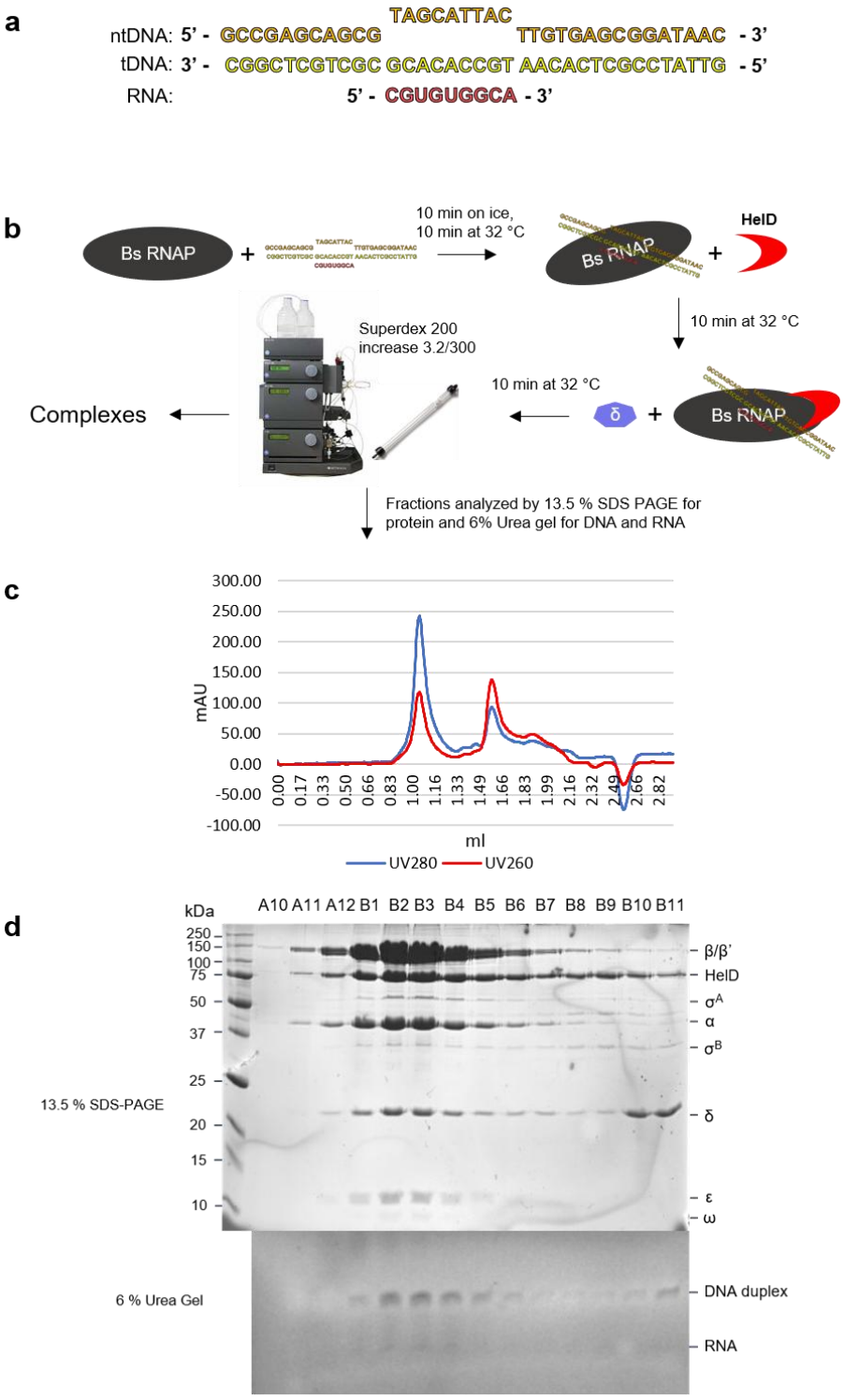


Figure 4.3 RNAP-δ-HelD complex assembly. a, The dsDNA-RNA hybrid, which contains an

artificial transcription bubble. b, RNAP-δ-HelD complex assembly strategy. c, Chromatogram of the SEC run of RNAP-δ-HelD complex in Superdex 200 increase 3.2/100 column. d, top: 13.5 % SDS-PAGE analysis of peak fractions from SEC. bottom: 6 %-Urea gel to analyze nuclear acids of fractions from SEC.

Since we did not know at which transcription step HelD engages the complex, we assembled the complex with a nucleic acid scaffold containing an artificial transcription bubble (Figure 4.3 a). The complex assembly strategy is shown in Figure 4.3 b.

Once fully assembled, the complex was loaded on a SEC an eluted in three peaks (Figure 4.3 c). The contents of each peak were further analyzed by 13.5 % SDS-PAGE for protein and 6 % urea gel for DNA and RNA.

As expected for its size and confirmed by fraction analysis, the complex elutes first from the column (Figure 4.3 d). As shown in Figure 4.3 d, fractions from this first peak (A11-B6) contain Bs RNAP, HelD, δ, σ^A and σ^B in SDS-PAGE, as well as DNA and RNA (as seen in the urea gel). Transcription initiation factors σ^A and σ^B are co-purified in Bs RNAP, which might form a holoenzyme with RNAP. As shown in the urea gel, the first peak contains DNA and RNA as well, this means we may also have an elongation complex in those fractions, but they are underrepresented. Interestingly, the ω subunit from Bs RNAP was also underrepresented compared to the original purified Bs RNAP. This might be due to minor losses of the ω subunit during complex assembly and SEC. Thus, the complex sample is likely consisting of several different complexes, including the σ^A-RNAP and σ^B-RNAP holoenzymes, the RNAP-DNA-RNA elongation complex and the RNAP-HelD-δ complex. The RNAP-HelD-δ complex is the major complex in the sample. However, it remains unclear if the complex contains σ^A, σ^B or nucleic acids in addition to RNAP, HelD and δ.

4.4 CryoEM data collection and processing of RNAP-δ-HelD

Although the obtained sample contained several different complexes, single-particle analysis (SPA) cryo-electron microscopy is a suitable method to determine the structure of RNAP-δ-HelD. Firstly, one can ideally identify different complexes in the sample by classification of particles. Moreover, the major content in the sample is RNAP-δ-HelD.

The SPA cryoEM sample was prepared with Dr. Tarek Hilal. The sample was prepared on plasma-treated Quantifoil R1/2 holey carbon grids right after the SEC run of the complex assembly. To overcome orientation problems, the sample was supplemented with 0.15 % (w/v) n-octylglucoside. The SPA cryoEM image data (Figure 4.4 a) were collected by Dr. Tarek Hilal on a FEI Titan Krios G3i (300 kV) with a Falcon 3EC camera.

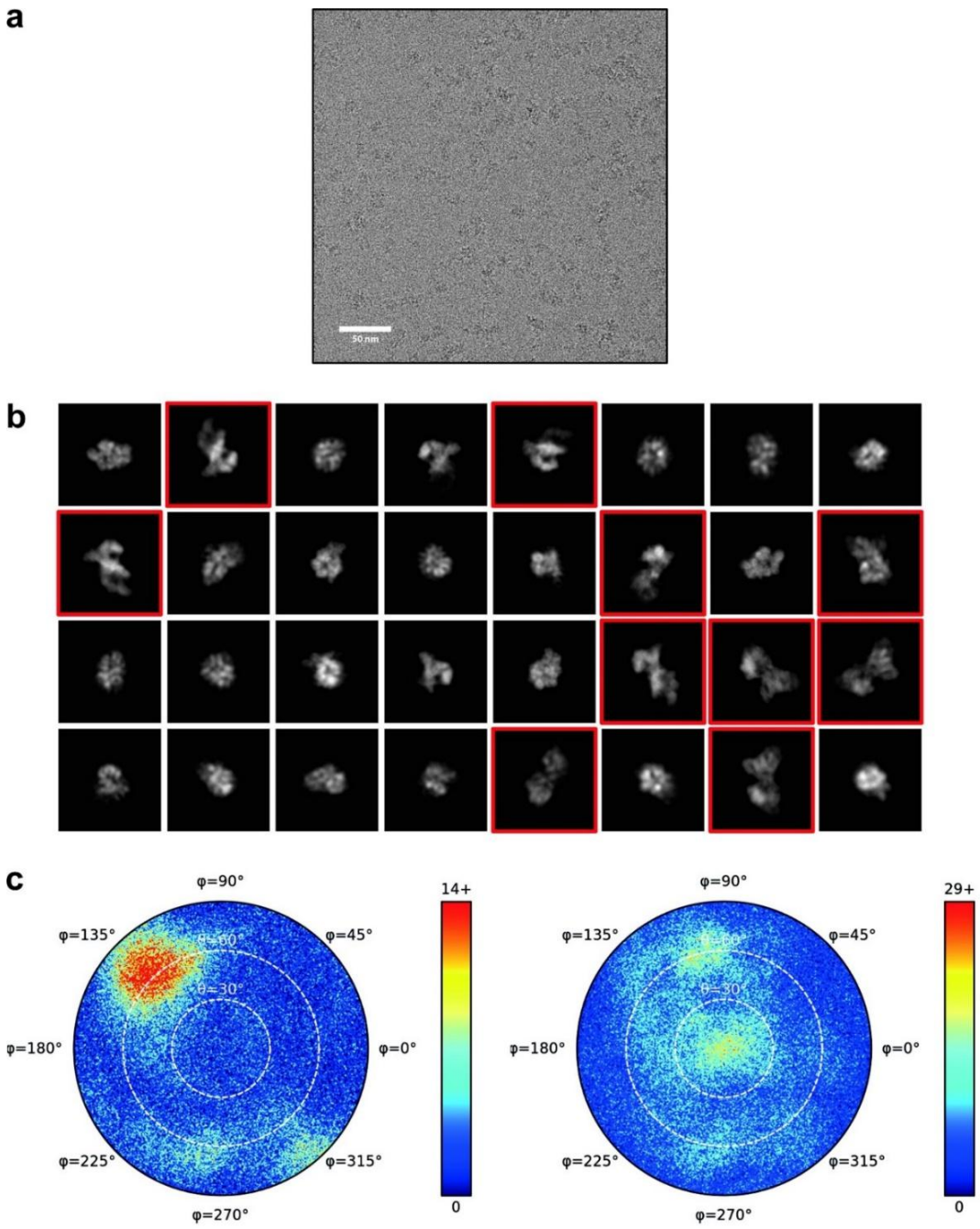


Figure 4.4: CryoEM data. a, Representative cryo electron micrograph of RNAP-δ-HelD particles. Scale bar, 50 nm. b, 2D class averages of RNAP-δ-HelD particles. Red boxes indicate classes representing RNAP-δ-HelD dimers. c, Polar plots of particle orientations in the monomeric (left) and dimeric (right) RNAP-δ-HelD datasets. Legends, color codes for particle numbers. (by Hilal, T.)

Dr. Tarek Hilal processed the cryoEM data and detected individual molecule particles in the images of RNAP-δ-HelD SPA cryoEM sample (Figure 4.4 a). These particles could subsequently be used for 3D structure reconstruction. 2D classification resulted in two types of the RNAP-δ-HelD complex, a monomeric and a dimeric one (Figure 4.4 b). Supplementation of the detergent n-octylglucoside solved the orientation problem for both the monomeric and the dimeric data sets (Figure 4.4 c).

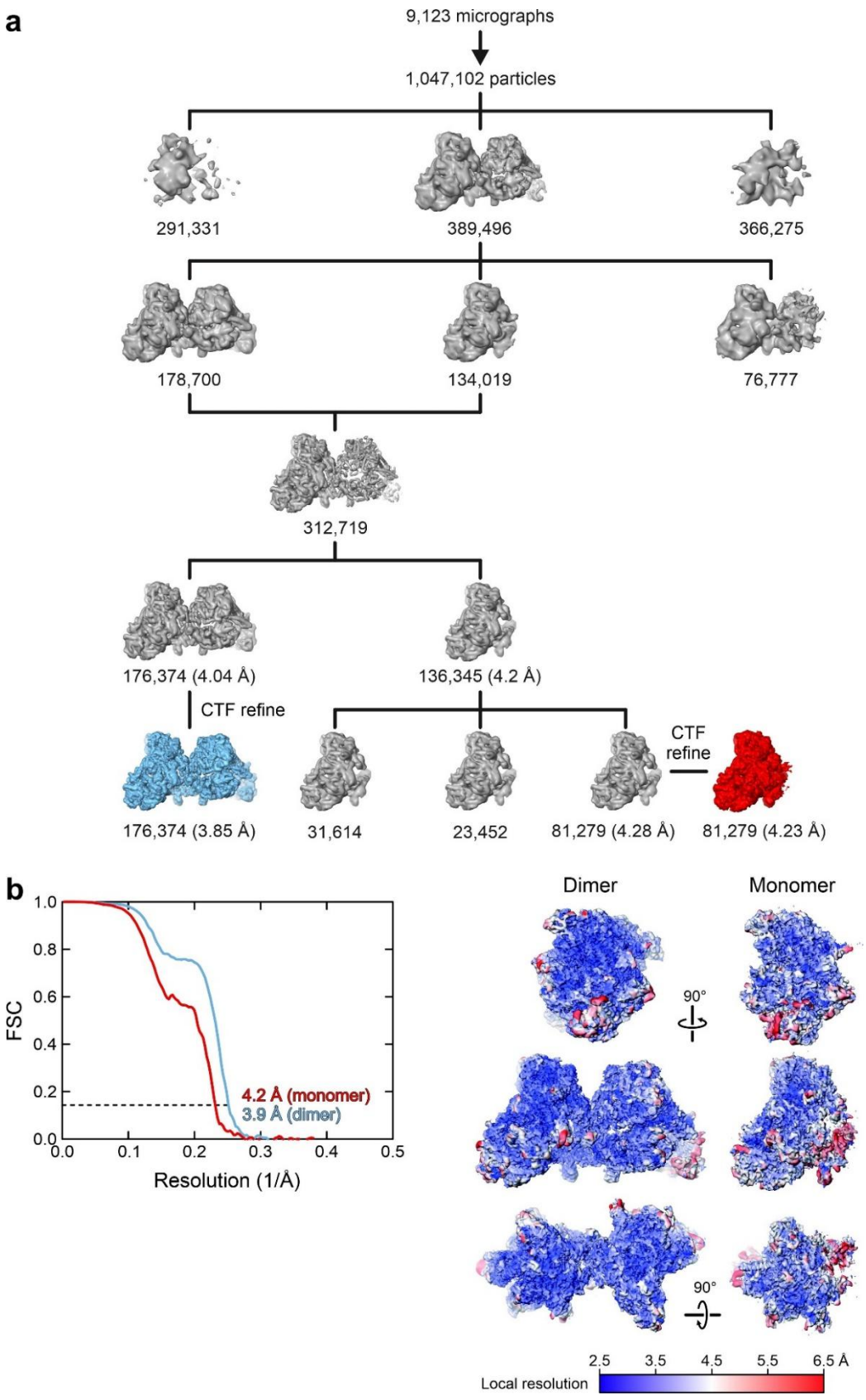


Figure 4.5: Hierarchical clustering analysis. a, Hierarchical clustering analysis. b, Left, Fourier shell correlation (FSC), indicating nominal resolutions of 4.2 Å and 3.9 Å for the overall cryoEM

maps of monomeric and dimeric RNAP-δ-HelD complexes, respectively, according to the FSC0.143 criterion. Right, orthogonal views of the cryoEM map for dimeric RNAP-δ-HelD, colored according to the local resolution in different regions. Legend, color code for local resolution. (by Hilal, T.)

To do multi-particle 3D refinement, about 1,000,000 particle images were selected from about 9,100 micrographs (Figure 4.5 a). After the refinement, we got one map for the monomeric RNAP-δ-HelD and one for the dimeric $(RNAP\text{-}\delta\text{-}HelD)_2$ complexes. Global resolutions ranged from 4.2 Å for the monomer and 3.9 Å for dimer (Figure 4.5 b). The local resolutions extended beyond these limits. These were also kindly done by Dr. Tarek Hilal.

4.5 RNAP-δ-HelD complex structure determination

RNAP-δ-HelD complex structure determination was done as described in the "Materials and Methods" section. With 3D structure reconstruction from SPA cryoEM, we obtained a monomeric and a dimeric RNAP-δ-HelD 3D electron density map. In both the monomeric and the dimeric complexes map, we observed well-defined density for the RNAP subunits α1/2 (N-terminal domains [NTDs]), β, β', δ, ε and HelD (Figure 4.6). We could, however, not see any density for the ω subunit or nucleic acids. The RNAP-δ-HelD complex model (Figure 4.6) was built together with Dr. Bernhard Loll and was based on the dimer map since it had a better resolution than the monomer one (Figure 4.5).

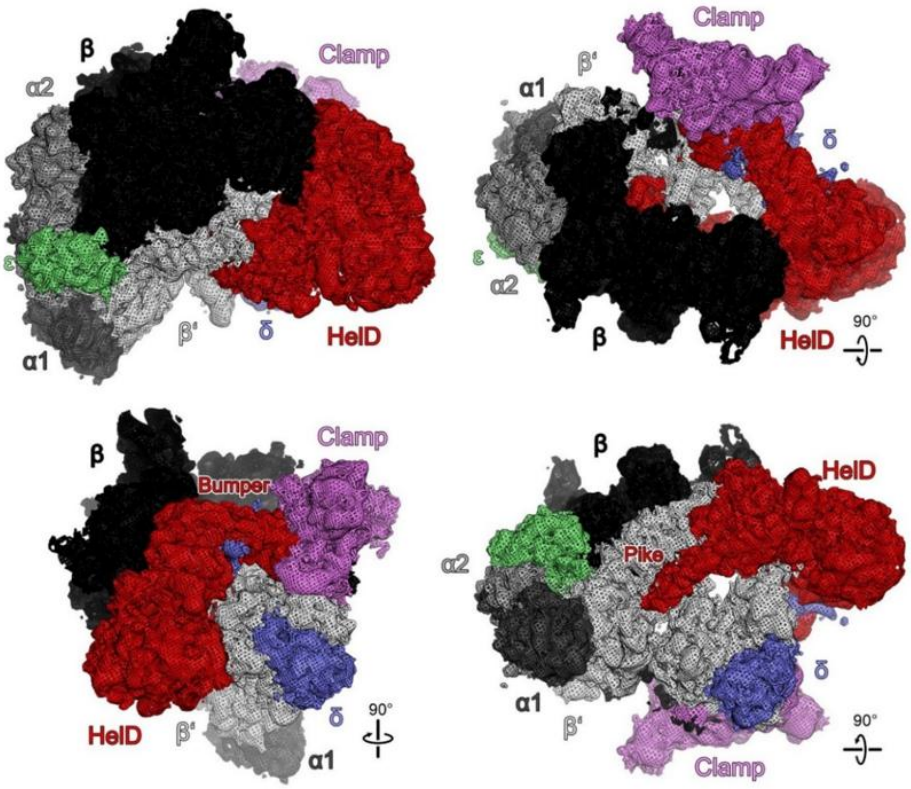


Figure 4.6: Orthogonal views of cryoEM density of the monomeric RNAP-δ-HelD complex contoured at the 7 σ level. (with Loll, B.)

Table 4.1 CryoEM data collection and refinement statistics.

Dataset	Monomeric RNAP-δ-HelD	Dimeric RNAP-δ-HelD
PDB ID	6ZCA	6ZFB
EMDB ID	EMD-11104	EMD-11105
Data collection		
Pixel size (Å/px)	0.832	
Defocus range (μm)	0.5 – 2.5	
Voltage (kV)	300	
Electron dose (e^-/Å^2)	40	
Number of frames	29	
Micrographs	9123	
Particle images	81279	176374
Refinement		
Global resolution $FCS_{0.143}$ (Å)	4.23	3.85
Local resolution range (Å)	2.6 – 7.0	2.8 – 6.5
CC mask	0.76	0.78
CC volume	0.75	0.77
Model composition		
Non-hydrogen atoms	28,724	57,521
Protein residues	3,640	7,313
Zn^{2+}/Mg^{2+} ions	1	2
Rmsd from ideal geometry		
Bond lengths (Å)	0.003	0.002
Bond angles (°)	0.535	0.559
Ramachandran plot (%)		
Favored	92.3	91.4
Allowed	7.7	8.4
Outliers	0.0	0.2
Model quality[a]		
Clash score	13.4	8.2
Rotamer outliers	0.0	4.0
Overall score	2.1	2.6

4.6 Organization of RNAP in an RNAP-δ-HelD complex

In the RNAP-δ-HelD complex (Figure 4.7), RNAP forms a conformation in which the main channel, where downstream DNA and the RNA:DNA hybrid are located in an EC,

is wide open. The distance between the β2 lobe (P242) and the β' clamp helices (N283) is about 52 Å. In addition, this wide open RNAP has an associated extension of the RNA exit tunnel by more than 17 Å (β flapR800 to β' lidD245). The α1/2 NTD dimer sits at the closed end of the open β/β' crab. (Domain names are shown in Table 1.1)

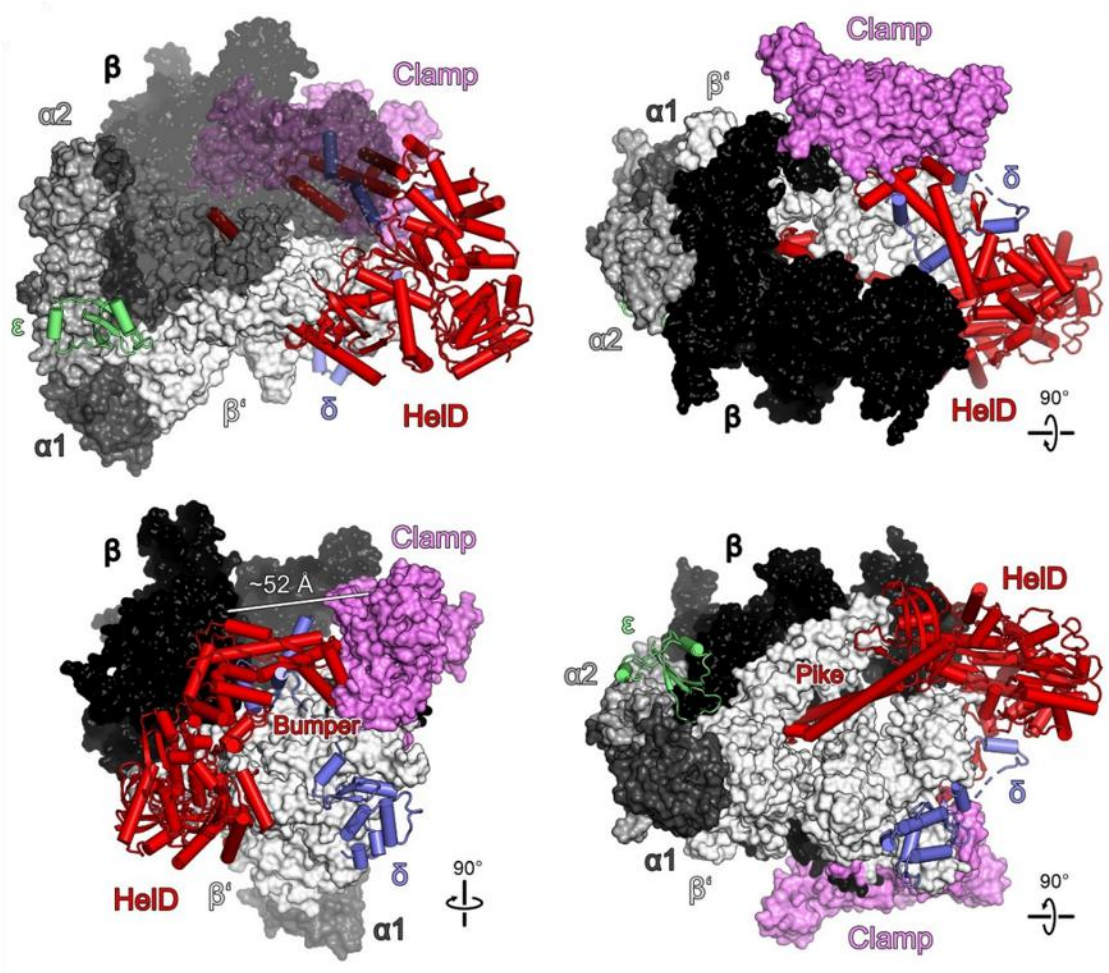


Figure 4.7: Overall architecture of a RNAP-δ-HelD complex. β surface is semi-transparent in the upper left panel. Rotation symbols in this and all figures indicate views relative to the upper left panel. Color coding in all figures, unless otherwise noted: α1, dark gray; α2, gray; β, black; β', light gray; β' clamp, violet; ε, lime green; δ, slate blue; HelD, red.

δ consists of a globular folded N-terminal domain (NTD; residues 1-90) and an intrinsically disordered acidic C-terminal region (CTR; residues 91-173) with a net negative charge of -47 (López de Saro, F.J. et al., 1995; Kubáň, V. et al., 2019). As mentioned earlier (Zachrdla, M. et al., 2017), the first ~70 residues of δ^{NTD} folded similar to the globular domain of the group 1 σ factors' σ1.1 region (Schwartz, E.C. et al., 2008). However, in contrast to the σ1.1 domain in an *E. coli* σ^{70} holoenzyme (Chen, J. et al., 2019), δ^{NTD} is not located in the main channel, but binds to the RNAP surface between the β' shelf and jaw (Figure 4.7). This is consistent with previous *in vivo* CLMS analysis (de Jong, L. et al., 2017).

The ε subunit resides in the cavity formed by the α1/2 NTDs, the C-terminal β clamp and residues 492-655 of β', which form part of the secondary channel (Figure 4.7). These arrangements stand in contrast to previous illustrations of ε on the β' jaw that are based on a low-resolution cryo EM analysis and structural similarity of ε with the phage T7 Gp2 (Keller, A.N. et al., 2014).

4.7 Structure of HelD in an RNAP-δ-HelD complex

HelD consists of four domains/regions: an N-terminal region (NTR; residues 4-187), two globular domains (D1a/D1b, residues 188-338/491-603; D2, residues 604-774), and an elongated helical protrusion in D1 ($HelD^{Bumper}$; residues 339-490; Figure 4.8 a). The NTR shows a remarkable similarity to GreA/B transcript cleavage factors, but with an elongated coiled-coil ($HelD^{Pike}$; residues 4-96; Figure 4.8 b). D1 and D2 are similar to NTPase/helicase domains of UvrD (Lee, J.Y. and Yang, W. 2006), only with a subdomain deleted from D2 and $HelD^{Bumper}$ inserted into D1 (Figure 4.8 c). $HelD^{Bumper}$ does not have any close structural similarity to other proteins in the Protein Data Bank (https://www.rcsb.org).

Table 4.2 Regions of HelD described in the text.

HelD		
NTD	-	1-200
D1a/D1b	-	201-338/491-603
Ins	-	339-490
D2	-	604-774

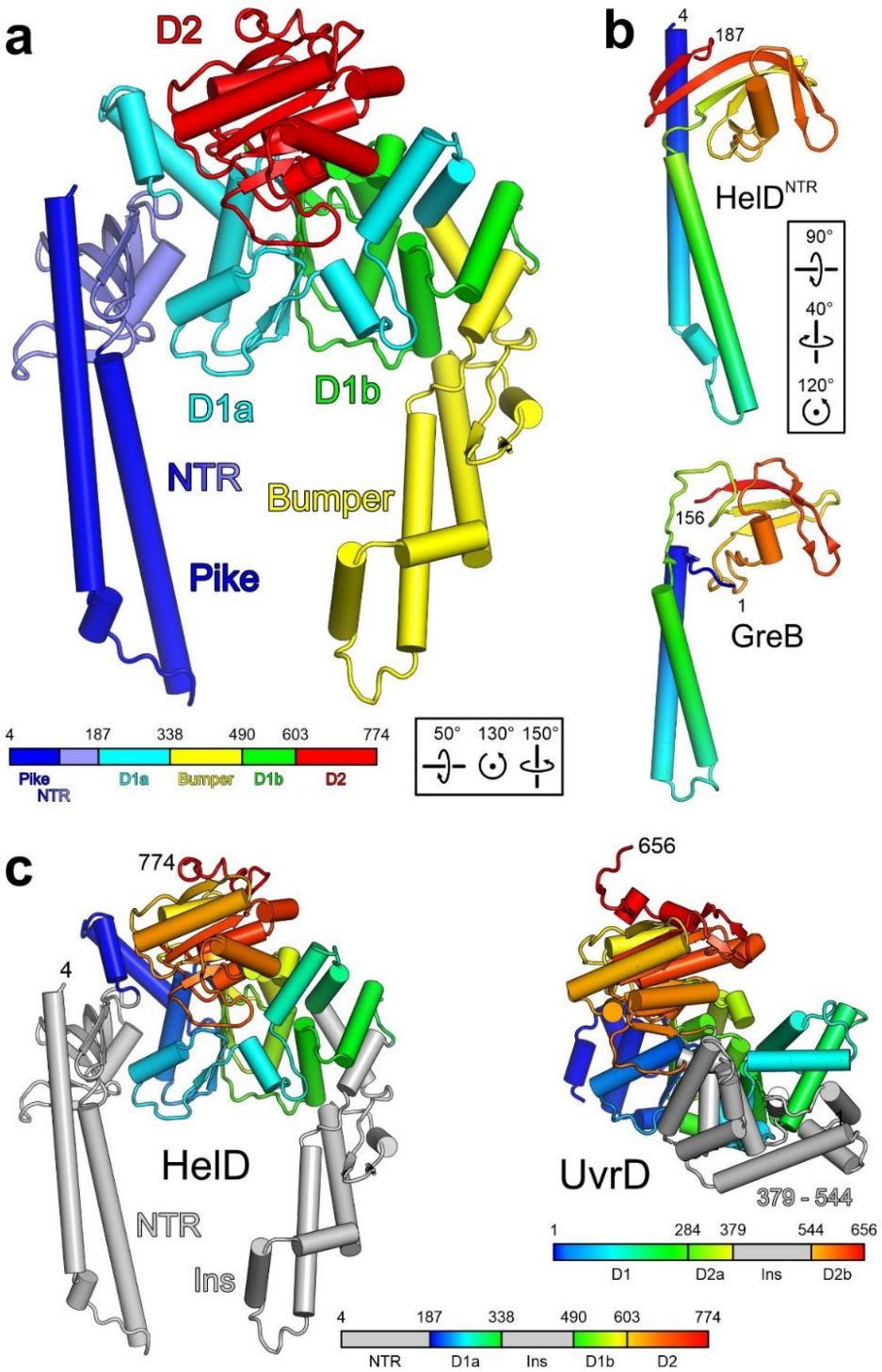


Figure 4.8: HelD architecture. a, Cartoon plot of HelD colored by domains (for color-coding see legend). Numbers refer to domain borders. b, Comparison of HelD^{NTR} to GreB (PDB ID 6RIN) reveals similar topology of the coiled-coils, which insert into the secondary channel, and the globular domains; in GreB, the latter is responsible for high-affinity binding to the RNAP β' rim helices. HelD^{NTR} and GreB are rainbow-colored (blue, N-termini; red, C-termini). Numbers refer to domain borders. c, Comparison of NTPase domains in HelD and in *E. coli* UvrD (PDB ID 2IS6). The D1-D2 regions are rainbow-colored (blue, N-termini; red, C-termini) as indicated in the legends. Neighboring and inserted regions (Ins), gray. Numbers refer to domain borders.

4.8 HelD and δ^{CTR} invades RNAP channels

HelD is reminiscent of a two-pronged fork that is inserted into RNAP. In perfect analogy to transcript cleavage factors (Abdelkareem, M. et al., 2019), one prong, $HelD^{Pike}$, is inserted deep into the secondary channel through which the substrate NTPs reach the active site of RNAP during elongation (Figure 4.9 a). D1/D2 grab around the β2 lobe and position the other prong, $HelD^{Bumper}$, in the main channel where it pushes against the β' clamp, pushing β and β' apart (Figure 4.7). In the course of HelD engaging RNAP, a large, combined surface area (~11,500 $Å^2$ total; ~8,000 $Å^2$ with β'; ~1,800 $Å^2$ with β; ~1,700 $Å^2$ with δ) is buried.

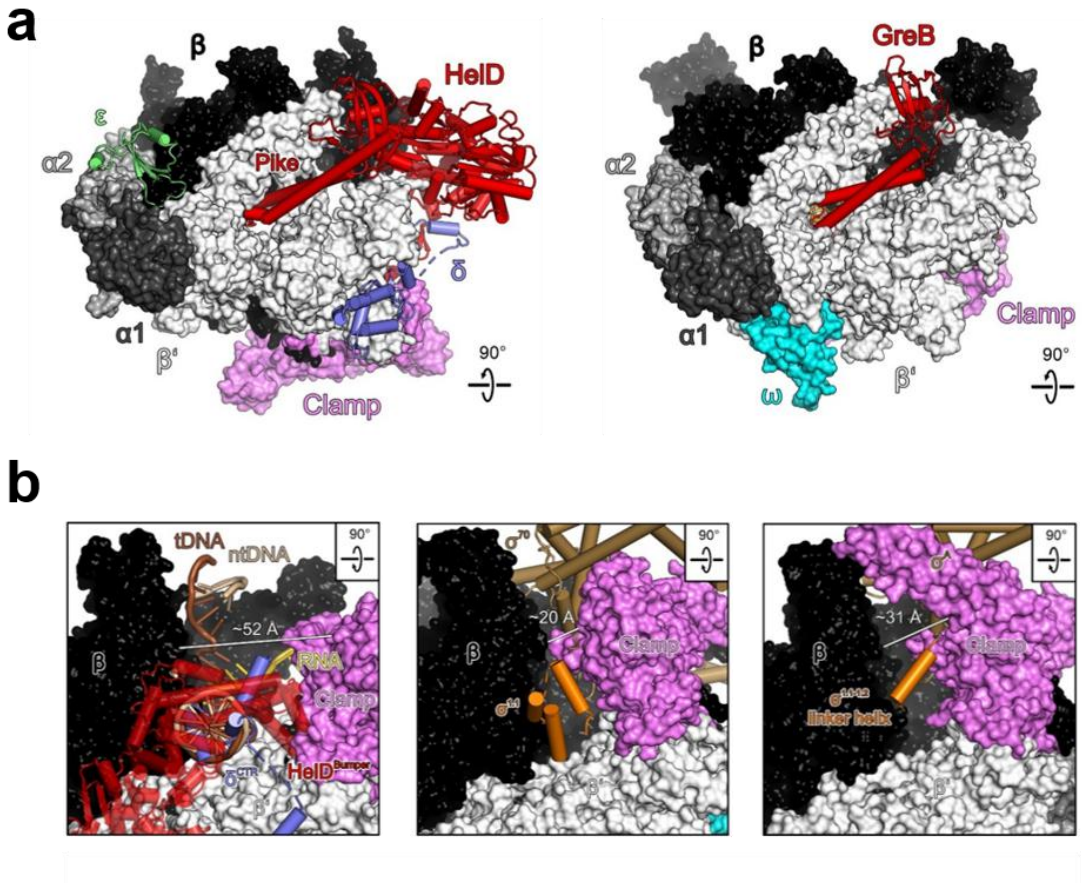


Figure 4.9 a, Comparison to an *E. coli* GreB-modified EC (PDB ID 6RIN), illustrating similar secondary channel invasion by coiled-coil elements in GreB and HelD. b, left: Nucleic acid scaffold from the *E. coli* EC (PDB ID 6ALH) transferred onto the RNAP-δ-HelD complex (HelD omitted) by superpositioning of the β subunits, showing competition of δ^{CTR} with the downstream DNA duplex in the main channel. middle: Comparison to an *E. coli* σ^{70} holoenzyme structure (PDB ID 6P1K), showing analogous positioning of δ^{CTR} and the σ1.1 globular domain in the main channel and the reduced channel width in the σ^{70} holoenzyme. σ^{70}, sand-colored; σ1.1, orange. right: Comparison to an *M. smegmatis* σ^{A} holoenzyme structure (PDB ID 6EYD), showing analogous positioning of δ^{CTR} and a σ1.1-1.2 linker helix in the main

channel and the reduced channel width in the σ^A holoenzyme. σ^A, sand-colored; σ1.1-1.2 linker helix, orange.

We see some cryoEM density patches around $HelD^{Bumper}$, which can be interpreted as parts of δ^{CTR} (Figure 4.7; Figure 4.9 b left). However, the quality of the local cryoEM density in that region is too poor to do a reliable modeling of the exact region of δ^{CTR} that binds to $HelD^{Bumper}$. By combining our results of HelD-δ^{CTR} interaction (Figure 4.2), δ^{CTR} might help to locate $HelD^{Bumper}$ in the main channel and helps HelD to push against the β' clamp (Figure 4.9 b left).

The preliminary modeling part of $HelD^{Bumper}$ and δ^{CTR} is located at the equivalent position of the spherical σ1.1 domain in the *E. coli* σ^{70} holoenzyme (Chen, J. et al., 2019) (Figure 4.9 middle) and a helix following the σ1.1 region in a *Mycobacterium smegmatis* σ^A holoenzyme (Kouba, T. et al., 2019) (Figure 4.9 right). Therefore, $HelD^{Bumper}$ and δ^{CTR} occupy the regions near the β subunit where the DNA is located in an EC (Figure 4.9 left).

4.9 RNAP-δ-HelD exhibits the most open main channel configuration

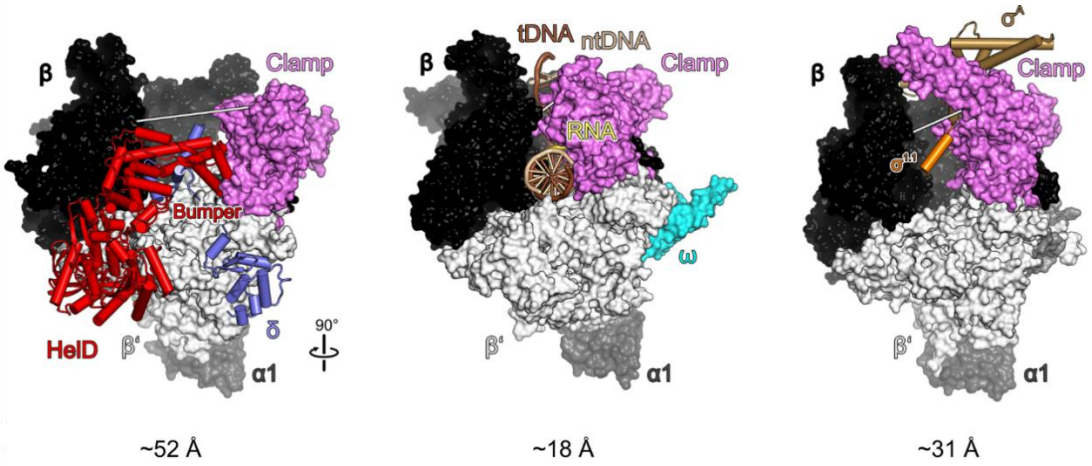


Figure 4.10 Comparison of main channel opening. left, *B. subtilis* RNAP-δ-HelD complex. middle, *E. coli* RNAP elongation complex (PDB ID 6ALF). right, *M. smegmatis* RNAP holoenzyme (PDB ID 6EYD).

Because of the joint actions of δ and HelD, *B. subtilis* RNAP-δ-HelD presents the most open main channel configuration that has been observed in RNAP complexes so far. It is ~52 Å wide and augmented by about 18 Å and 31 Å relative to the *E. coli* EC and *M. smegmatis* σ^A holoenzymes (Kang, J.Y. et al., 2017; Kouba, T. et al., 2019), respectively (Figure 4.10). This wide-open configuration may help to release the DNA duplex or nascent RNA from RNAP.

4.10 HelD^{Pike} dismantles the RNAP active site and competes with RNA

After penetrating the secondary channel, HelD^{Pike} locally destroys the β' bridge helix (BH; between residues 780 and 787) and locks the β' trigger loop (TL; Figure 4.11 a b); key elements that rearrange for nucleotide addition during elongation (Belogurov, G.A. and Artsimovitch, I. 2019). Although HelD^{Pike} has negatively charged side chains (D56, D57, E60) at its tip, these residues will not reshape the active site as observed in GreB-RNAP complex (Abdelkareem, M. et al., 2019). Instead, the tip is ploughed over the active center, thus dismantling it. When the β C-terminal clamp is pushed away from the nucleic acid, the β switch region 3 (Sw3) of the hybrid in the EC becomes disordered and the active site loop (ASL) is rearranged, thereby losing the catalytic Mg^{2+} ion (Figure 4.11 ab)

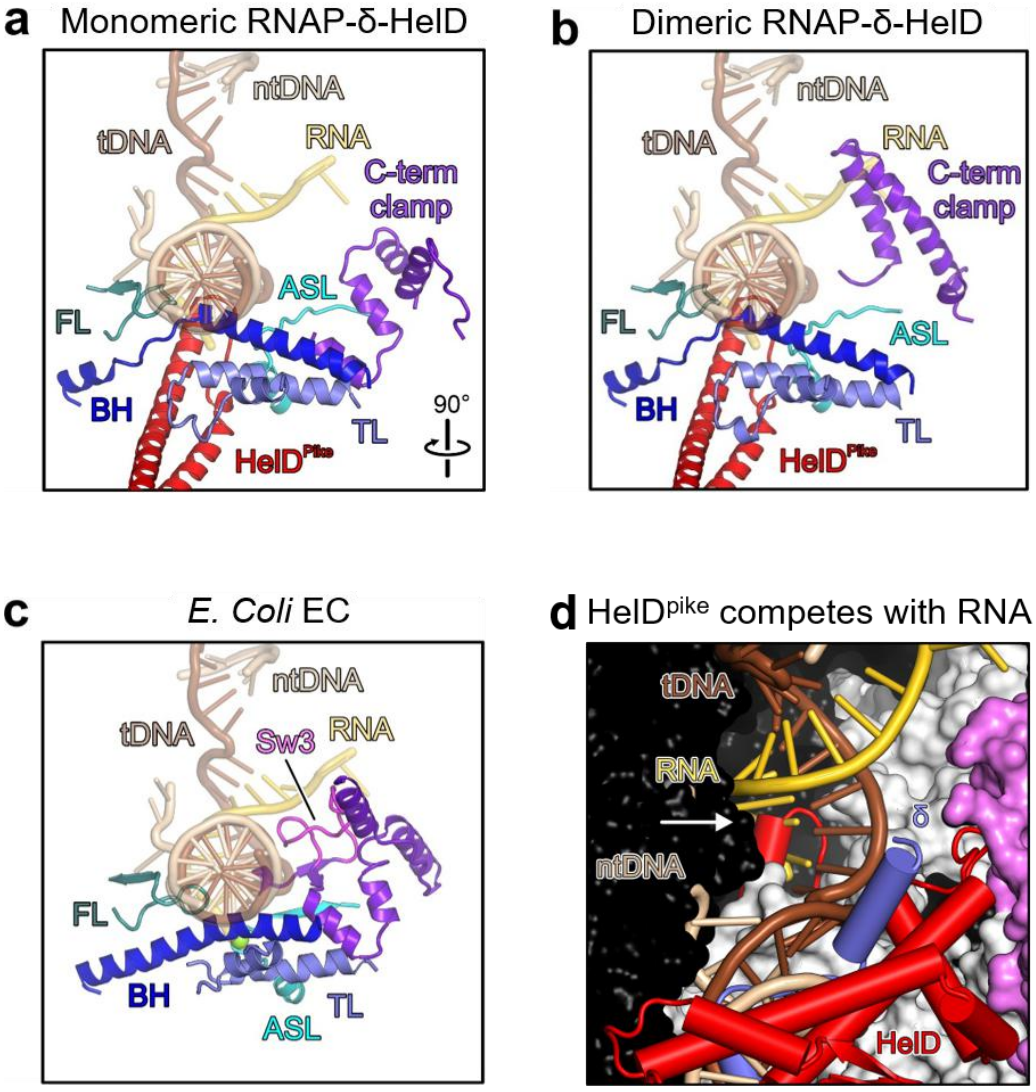


Figure 4.11 Active site dismantling and HelD^{pike} competes with RNA binding. a, b, RNAP active

site environments in monomeric RNAP-δ-HelD (a) and dimeric (b) RNAP-δ-HelD, showing HelD-mediated active site dismantling. Nucleic acids of an *E. coli* EC (PDB ID 6ALH) were transferred to RNAP-δ-HelD by superposition of the β subunits. β elements: FL (fork loop), teal; C-term clamp (C-terminal clamp), purple; Sw3, magenta. β' elements: ASL, cyan; BH (bridge helix), blue; TL (trigger loop), slate blue. The catalytic Mg^{2+} ion (green sphere) is lost from RNAP-δ-HelD upon $HelD^{Pike}$ invasion. c, Comparison to the RNAP active site environment in an *E. coli* EC (PDB ID 6ALH). d, Close-up view of RNAP active site region in RNAP-δ-HelD, with a nucleic acid scaffold from the *E. coli* EC (PDB ID 6ALH) transferred onto the RNAP-δ-HelD complex by superpositioning of the β subunits, illustrating direct competition of the $HelD^{NTR}$ coiled-coil tip with RNA (white arrow).

RNAP-RNA binary complexes have catalytic activities, suggesting RNA resides in the active site cavity (Altmann, C.R. et al., 1994). Compared with an *E. coli* EC (Kang, J.Y. et al., 2017), the $HelD^{Pike}$ tip directly competes with the RNA in the hybrid for binding to RNAP (Figure 4.11 d) and may additionally repel RNA through its negatively charged residues. Therefore, $HelD^{Pike}$ rearranges the active site regions and competes spatially with all nearby bound RNAs. The opening of the RNA exit tunnel through $HelD^{Bumper}$ will promote the release of RNA.

4.11 $HelD^{Bumper}$ and δ displace nucleic acids

Obviously, the binding of $HelD^{Bumper}$ and $δ^{CTR}$ in the main channel is incompatible with the DNA occupying this site (Figure 4.9 b left). Previous studies have shown that δ or excess $δ^{CTR}$ alone can release RNA or DNA from RNAP (López de Saro, F.J. et al., 1995). To further describe the contributions of δ and HelD to nucleic acid replacement, we performed band shift assays.

We first bound RNAP to nucleic acids and then added δ and/or HelD. We first tested the release of DNA with an artificial bubble, which when bound to RNAP mimics a situation what happens after many intrinsic termination events (Harden, T.T. et al., 2020; Belogurov, G.A. and Artsimovitch, I. 2019; Artsimovitch, I. and Landick, R. 2002). HelD displaced approximately 25 % of DNA from $RNAP^{ΔδΔHelD}$, while δ caused approximately 80 % displacement without HelD (Figure 4.12, lanes 4-6). In the presence of a stoichiometric amount of HelD, increasing amounts of δ titrated to DNA-bound $RNAP^{ΔδΔHelD}$ resulted in a gradual decrease of bound DNA. When an equimolar amount of δ relative to $RNAP^{ΔδΔHelD}$-HelD was added, almost all DNA was displaced (Figure 4.12, lanes 7-13). By adding equal molar amounts of HelD and $δ^{NTD}$, only about 50 % of the DNA was displaced (Figure 4.12, lane 14).

Next, we tested the ability of δ/HelD to displace RNAP from DNA with an artificial bubble and complementary RNA, mimicking δ/HelD-mediated recovery of RNAPs that became stuck during elongation without releasing RNA. A picture similar to that of only DNA displacement appeared; however, due to the RNA-mediated stabilizing effect of DNA on RNAP, HelD and δ individually or HelD/$δ^{NTD}$ released less RNAP

than in the case with only DNA bound. Moreover, a higher δ concentration is required in the presence of HelD to complete a full nucleic acid displacement (Figure 4.12, lanes 15-28). It is worth noting that HelD/δ-mediated DNA or DNA/RNA displacement did not require the addition of ATP. Together, these results explain why a nucleic acid scaffold cannot associate with the RNAP-δ-HelD complex during the preparation of cryoEM samples. Furthermore, they emphasize the importance of δ in nucleic acid displacement, indicating that HelD is necessary to achieve complete nucleic acid release.

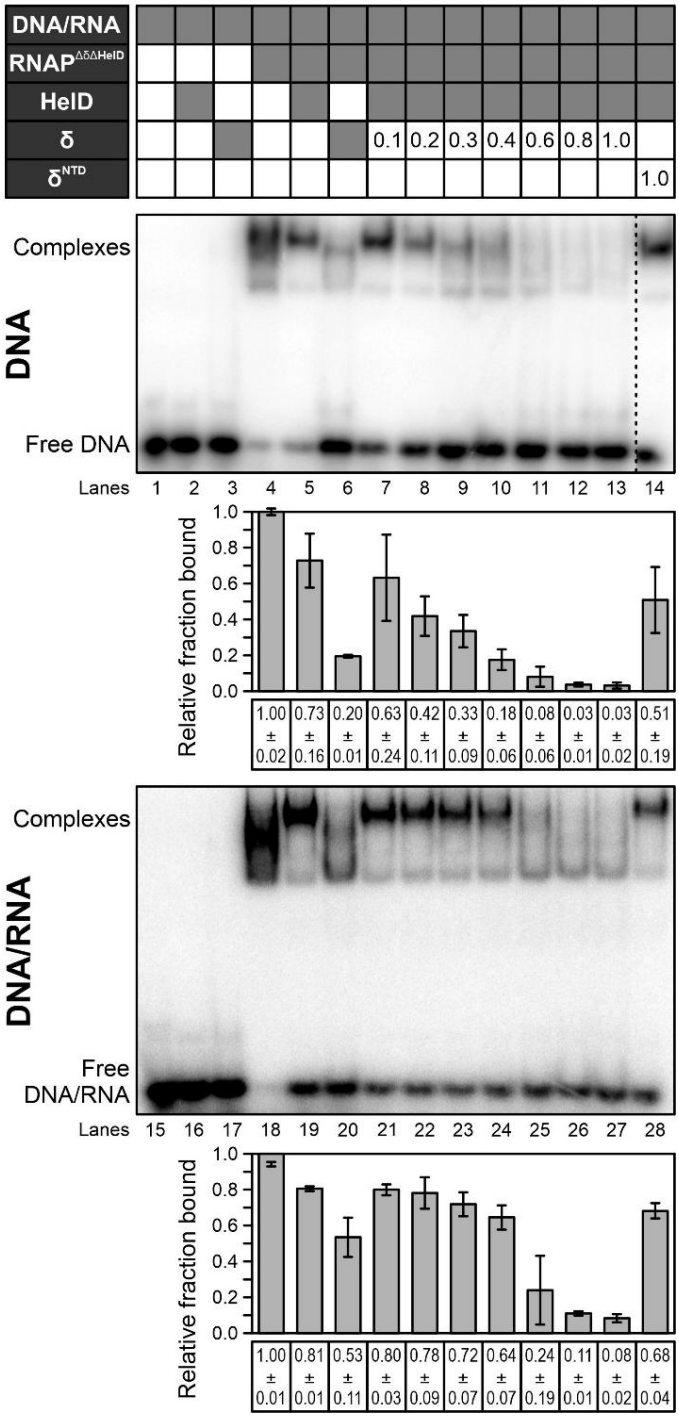


Figure 4.12 EMSA monitoring displacement of DNA (lanes 1-14) or DNA/RNA (lanes 15-28)

from RNAP by HelD, δ or combinations. Top scheme, samples analyzed; gray boxes, respective component added (proteins in equimolar amounts to $RNAP^{\Delta\delta\Delta HelD}$). Numbers, molar ratios of δ or δ^{NTD} relative to $RNAP^{\Delta\delta\Delta HelD}$ added. Panels labeled “DNA” or “DNA/RNA”, native PAGE analyses. All lanes are from the same gel, some lanes for the DNA-only gel were removed for display purposes (dashed line). Bar graphs, quantification of the data shown in the middle panels. Values represent means of DNA bound relative to $RNAP^{\Delta\delta\Delta HelD}$ alone ± SD for three independent experiments.

4.12 HelD-δ induced RNAP structural rearrangements confirmed by cross-linking mass spectrometry

Since our RNAP-δ-HelD complex exhibits the most wide-open main channel confirmation that has been observed in RNAP so far, we tried to confirm molecular contacts and HelD-triggered structural rearrangements by cross-linking coupled to mass-spectrometry.

In order to map molecular neighborhoods in the RNAP-δ-HelD complex, $RNAP^{\Delta\delta\Delta HelD}$ and recombinant δ and HelD were used to assemble $RNAP^{\Delta\delta\Delta HelD}$-δ, $RNAP^{\Delta\delta\Delta HelD}$-HelD and $RNAP^{\Delta\delta\Delta HelD}$-δ-HelD. The complexes were subsequently analyzed *via* CLMS with the heterobifunctional, photoactivatable crosslinker sulfosuccinimidyl 4,4′-azipentanoate (sulfo-SDA; Figure 4.13 a, b; Table 4.3 and Appendix 7.2). I prepared samples for CLMS together with Zhuo A. Chen, and Zhuo A. Chen performed the MS measurements. Matching the δ^{NTD} binding site deduced by cryoEM, a short stretch of δ residues crosslinked to the β’ jaw in both $RNAP^{\Delta\delta\Delta HelD}$-δ ($\delta^{Y82,P83,Y85}$-$\beta'^{K1032}$) and $RNAP^{\Delta\delta\Delta HelD}$-δ-HelD ($\delta^{Y83,Y85,L87,E90}$-$\beta'^{K1032}$). In the main channel, along the area connecting the main channel and the secondary channels, and in the active site area, multiple crosslinks of HelD were identified for $RNAP^{\Delta\delta\Delta HelD}$-HelD and $RNAP^{\Delta\delta\Delta HelD}$-δ-HelD complexes, which are connected to our cryoEM structures in excellent agreement (Figure 4.13 g).

$RNAP^{\Delta\delta\Delta HelD}$, $RNAP^{\Delta\delta\Delta HelD}$-δ and $RNAP^{\Delta\delta\Delta HelD}$-HelD produced significantly more crosslinks than $RNAP^{\Delta\delta\Delta HelD}$-δ-HelD, and especially compared with the RNAP-δ-HelD structure, more over-length crosslinks were formed (Figure 4.13 c, d). In addition, compared with complexes containing HelD, the proportion of crosslinks corresponding to over-length crosslinks was greatly increased in $RNAP^{\Delta\delta\Delta HelD}$ and $RNAP^{\Delta\delta\Delta HelD}$-δ (Figure 4.13 c, d). The reduced total number of crosslinks indicates a reduction in conformations explored by RNAP upon δ or HelD binding, especially when both factors are present. The reduced total number and fraction of over-length crosslinks in the presence of HelD indicates a conformation closer to our RNAP-δ-HelD cryoEM structure. A specific set of crosslinks between the β1/2 lobes (residues 146-248) and the β’ shelf and jaw (residues 794-1141) represent conformations in which β and β’ are close to each other across the main channel unless both δ and HelD are bound to RNAP (Figure 4.13 e, f). Together, our results

illustrate that HelD interacts with the main and the secondary channels of RNAP and that a stable main channel opening depends on the presence of both δ and HelD.

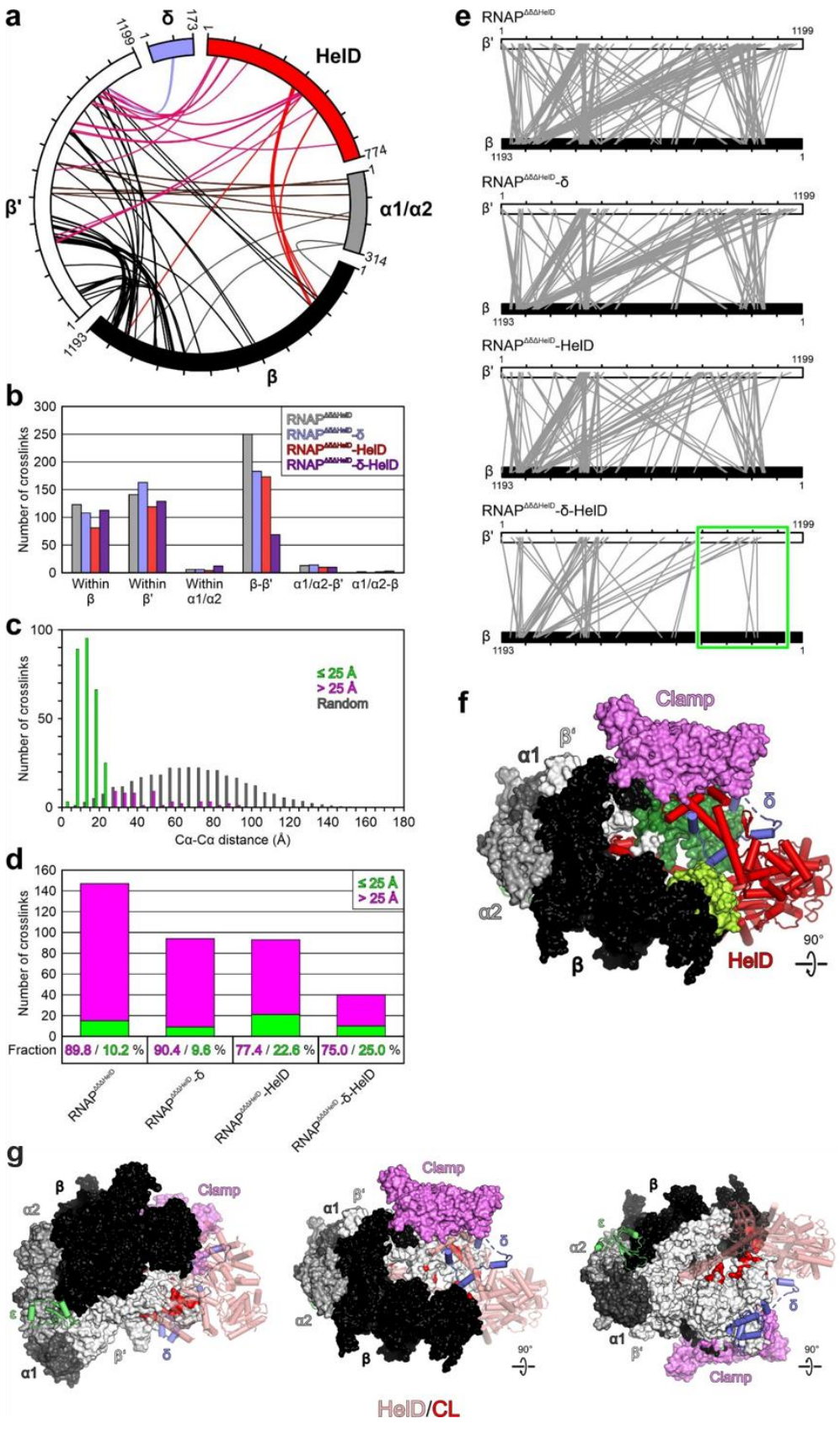


Figure 4.13 Structure probing by CLMS. a, Map of hetero-protein crosslinks observed in RNAP$^{\Delta\delta\Delta HelD}$-δ-HelD complex. b, Crosslinks identified in RNAP$^{\Delta\delta\Delta HelD}$, RNAP$^{\Delta\delta\Delta HelD}$-δ, RNAP$^{\Delta\delta\Delta HelD}$-HelD and RNAP$^{\Delta\delta\Delta HelD}$-δ-HelD. Binding of both δ and HelD leads to strongly reduced crosslinking between β and β'. c, Distribution of Cα-Cα distances between crosslinked residue pairs in reference to the RNAP-δ-HelD structure. Crosslinks with Cα-Cα distances within 25 Å, the theoretical crosslinking limit of sulfo-SDA, green; crosslinks with Cα-Cα distances > 25 Å, magenta; distance distribution of random residue pairs in the RNAP-δ-HelD structure, gray. d, Numbers of crosslinks (bars) between β and β' identified from the four crosslinked complexes, and fractions of over-length crosslinks (percentages at the bottom). Crosslinks are color-coded as in (b). In the RNAP$^{\Delta\delta\Delta HelD}$-δ-HelD complex, a significantly reduced number of β-β' over-length crosslinks (in reference to the RNAP-δ-HelD structure) compared to the RNAP$^{\Delta\delta\Delta HelD}$, RNAP$^{\Delta\delta\Delta HelD}$-δ and RNAP$^{\Delta\delta\Delta HelD}$-HelD complexes suggests that δ and HelD cooperate to stabilize an open conformation of RNAP. e, Comparison of β-β' crosslinks observed with RNAP$^{\Delta\delta\Delta HelD}$, RNAP$^{\Delta\delta\Delta HelD}$-δ, RNAP$^{\Delta\delta\Delta HelD}$-HelD and RNAP$^{\Delta\delta\Delta HelD}$-δ-HelD. Green boxed region, crosslinks between the β1/2 lobes (residues 146-248) and the β' shelf and jaw (residues 794-1141) observed in the first three complexes but almost absent in RNAP$^{\Delta\delta\Delta HelD}$-δ-HelD. f, Structure of the RNAP-δ-HelD complex highlighting the β1/2 lobes (residues 146-248; lemon green) and β' shelf and jaw (residues 794-1141; forest green), which largely lack crosslinks in the presence of δ and HelD (green box in e). g, Mapping of CLs to the structure. Mapping of HelD crosslinks to the surface of RNAP. HelD, semi-transparent, light red; crosslinked RNAP residues (CL), red. (by Chen, Z.A.)

Table 4.3 HelD and δ inter-molecular crosslinks in RNAP$^{\Delta\delta\Delta HelD}$-δ-HelD.

Protein/region 1	Residue 1	Protein/region 2	Residue 2	Highest score	Cα-Cα (Å)
HelD	442	β' rudder	312	13.34	13.7
HelD	37	β' trigger loop	948	16.376	12.2
HelD	37	β' trigger loop	949	15.601	9.3
HelD	37	β' trigger loop	939	14.106	19.6
HelD	37	β' bridge helix	831	8.929	13.3
HelD	37	β' bridge helix	832	10.791	13.9
HelD	37	β' bridge helix	834	13.462	11.2
HelD	37	β' bridge helix	837	9.694	15.4
HelD	433	β' bridge helix	800	12.841	14.9
HelD	437	β' bridge helix	793	14.772	17.1
HelD	45	β' 2° channel	617	10.918	14.2
HelD	443	β' clamp	314	11.393	7.9
HelD	443	β' clamp	317	7.78	8.5
HelD	392	β 2-lobe	177	11.621	16.9
HelD	393	β 2-lobe	177	14.001	14.4
HelD	397	β 2-lobe	190	11.908	15.4
HelD	386	β 2-lobe	223	16.669	9.3

HelD	388	β 2-lobe	223	9.172	11.2
HelD	392	β 2-lobe	223	15.132	12.6
HelD	392	β 2-lobe	223	15.132	12.6
HelD	393	β 2-lobe	223	15.204	11.2
HelD	396	β 2-lobe	223	11.95	15.8
HelD	397	β 2-lobe	223	14.267	15.8
HelD	399	β 2-lobe	223	13.703	20.4
HelD	537	β 2-lobe	216	11.78	13.6
HelD	535	β 2-lobe	223	11.126	14.7
HelD	536	β 2-lobe	223	13.548	13.1
HelD	97	β' jaw	983	13.815	16.9
HelD	97	β' jaw	1006	12.288	15.5
HelD	97	β' jaw	1007	10.609	14.8
HelD	97	β' jaw	1010	11.021	12.7
HelD	93	β' jaw	1011	8.759	12.4
HelD	175	β' jaw	1012	11.727	15.5
HelD	425	β' jaw	1031	13.164	23.5
HelD	424	β' jaw	1032	11.964	21.0
HelD	426	β' jaw	1032	16.241	18.9
HelD	427	β' jaw	1032	18.178	21.4
HelD	425	β' jaw	1033	10.684	24.6
δ	83	β' jaw	1032	12.628	19.3
δ	85	β' jaw	1032	15.907	16.6
δ	87	β' jaw	1032	11.341	16.0
δ	90	β' jaw	1032	13.574	19.4

Color intensity scales with proximity of crosslinked RNAP elements to the active site.

4.13 ATP-dependent HelD release

Since HelD completely disables RNAP (Figure 4.11 a, b), it must be released to allow transcription to resume. In order to study how HelD is released from the RNAP-δ-HelD complex, we performed several interaction analytical SEC. We firstly tested whether the transcription initiation factor can lead to a release of HelD. As shown in Figure 4.14 a, σ^A did not release HelD from the RNAP-δ-HelD complex in SEC. Because HelD has an ATPase domain, we thereby investigated the role of ATP in HelD-RNAP interaction. By comparing HelD to UvrD bound to DNA and ADP-Mg_2F_3 (Lee, J.Y. and Yang, W. 2006), we found that the D1/D2 conformation of RNAP-bound HelD is not capable to bind ATP (Figure 4.14 b). As detected via SAXS (Kovaľ, T. et al., 2019), ATP-binding can induce conformational changes in HelD. We therefore surmised that ATP-binding may lead to a release of HelD from RNAP. To verify this hypothesis, we performed HelD release assays as described in the “Materials and Methods” section. The results show that ATPγS, AMPPNP and, to a

somewhat lesser extent, ATP lead to the release of HelD from RNAP-δ-HelD during SEC, while ADP or AMP have minor effects (Figure 4.14 c). HelD has an intrinsic ATPase activity that is independent of RNAP (Wiedermannová, J. et al., 2014). Thus, AMPPNP and ATPγS mimic the constantly high ATP supply conditions, whereas ATP is probably hydrolyzed and separated from RNAP/HelD during SEC, reducing its effect. However, the addition of ATP or analogs did not lead to a displacement of δ from RNAP (Figure 4.14 a).

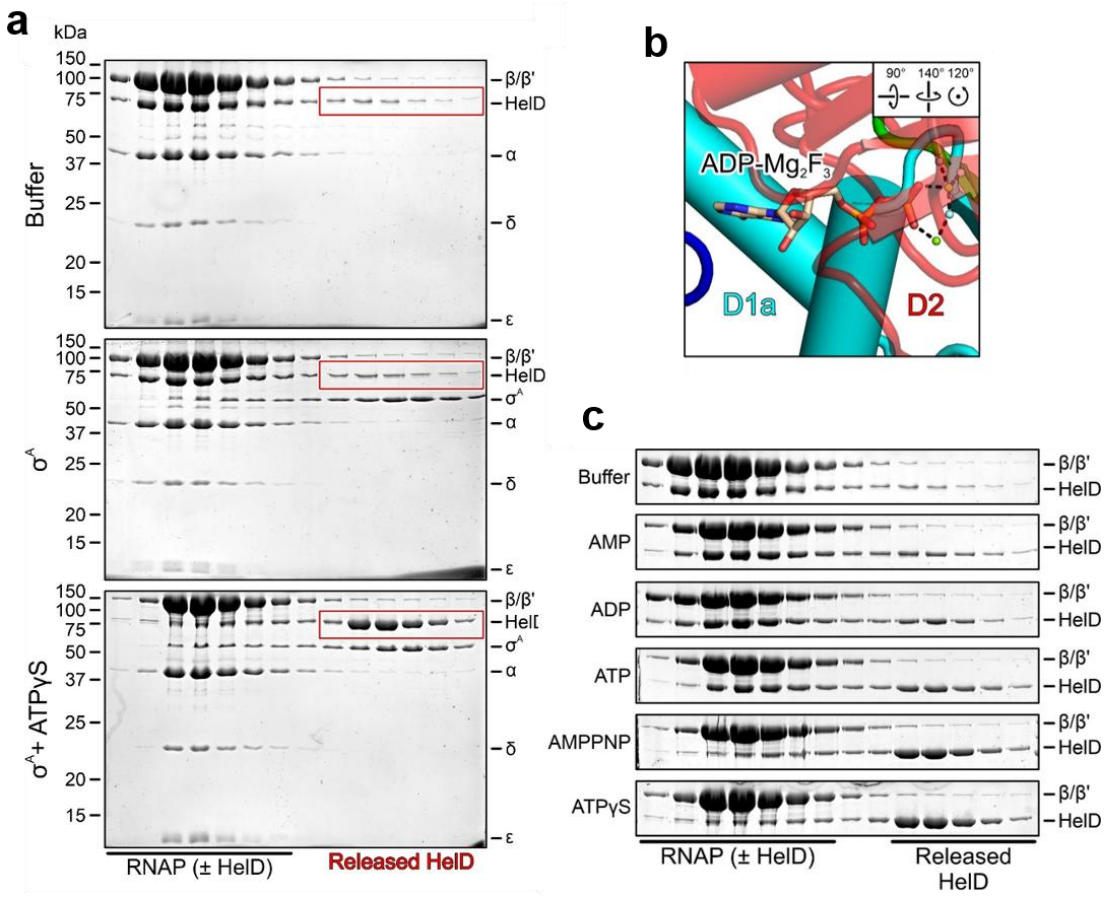


Figure 4.14 ATP-dependent HelD release. a, SDS-PAGE analysis of SEC runs after treating RNAP-δ-HelD with buffer, σ^A or σ^A/ATPγS, as indicated on the left. b, Close-up view of the ATP binding site of HelD, with ADP-Mg_2F_3 from a UvrD complex (PDB ID 2IS6) transferred by superpositioning of the UvrD NTPase domains on HelD, illustrating clashes with the nucleotide. ADP-Mg_2F_3 shown as sticks and colored by atom type; carbon, beige; nitrogen, blue, oxygen, red, phosphorus, orange; magnesium ions, green; fluoride ions, light blue. c, 12.5 % SDS-PAGE analysis of SEC runs after treating RNAP-δ-HelD with buffer or the nucleotides indicated on the left.

4.14 Dimeric (RNAP-δ-HelD)$_2$

About two thirds of our cryoEM particle images conformed to dimeric (RNAP-δ-HelD)$_2$ complexes (Figure 4.15 a; section 7.6 Supplementary Data S3). The protomers of the

dimeric RNAP-δ-HelD complex assembly is very similar to the monomeric one (root-mean-square deviation of 1.2-1.3 Å for 23,360-23,971 pairs of Cα atoms), but elements at the active site of RNAP are further remodeled in the dimers (Figure 4.11 a, b). The HelD-repositioned clamp forms an important contact region in the dimer, which also involves the initiation/elongation factor-binding β flap tip (FT; Figure 4.15 a). The dimeric $(RNAP\text{-}\delta\text{-}HelD)_2$ complex is very similar to the hibernating dimeric eukaryotic RNAP I (Pol I) (Fernández-Tornero, C. et al., 2013; Fernández-Tornero, C. 2018; Neyer, S. et al., 2016), where similar regions contribute to the dimer interfaces (Figure 4.15 b). Like in $(RNAP\text{-}\delta\text{-}HelD)_2$, each protomer of the hibernating RNAP I dimer shows a wide-open DNA-binding cleft, partially unfolded bridge helix and a DNA-mimicking loop stably bound inside the cleft (Fernández-Tornero, C. et al., 2013; Fernández-Tornero, C. 2018; Neyer, S. et al., 2016), similar to δ^{CTR}. Moreover, the A12.2 C-terminal domain of RNAP I is located inside the secondary channel (Fernández-Tornero, C. 2018). These observations indicate that, like the RNAP I dimer, the dimeric RNAP-δ-HelD might represent a dormant state.

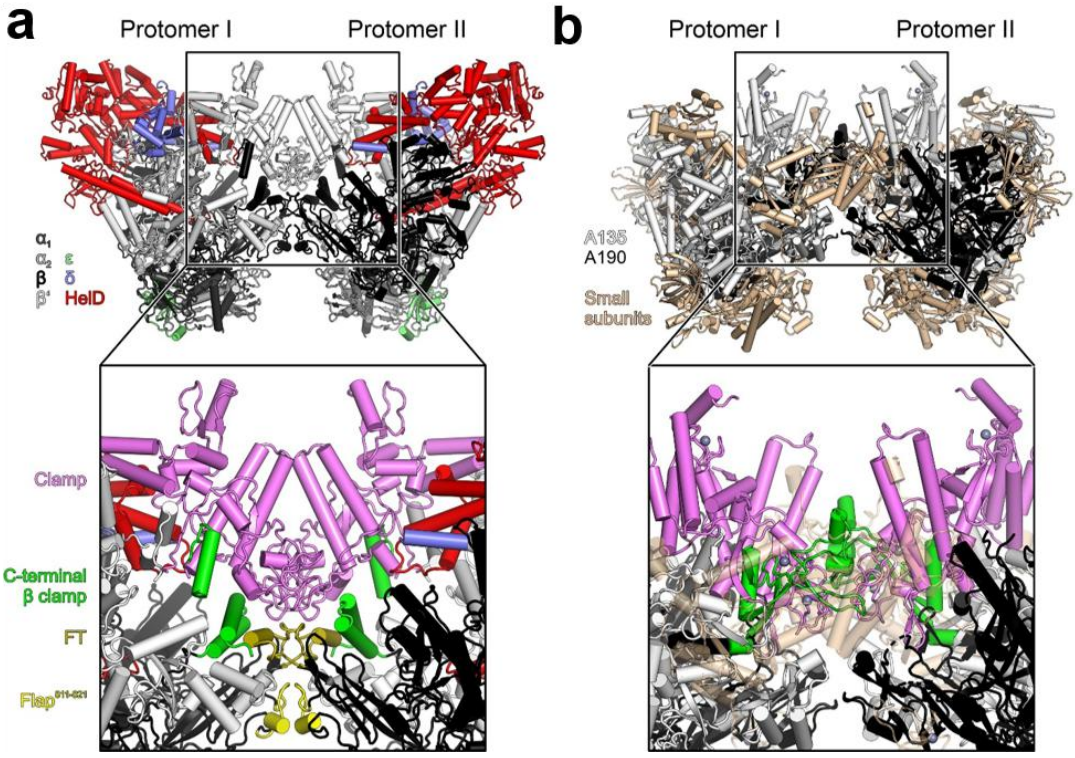
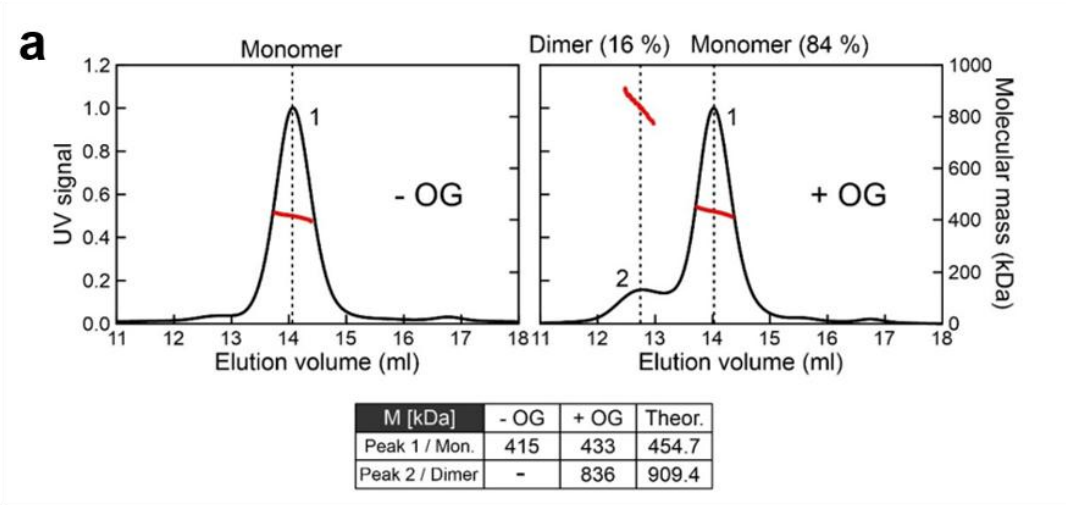


Figure 4.15 Dimeric $(RNAP\text{-}\delta\text{-}HelD)_2$. a, Structure of dimeric RNAP-δ-HelD. Inset, close-up view on the dimer interface. The two protomers interact via the elements highlighted in colors; β' clamp, violet; C-terminal β clamp, green; β flap tip (FT), olive; residue 811-821 of the β flap ($Flap^{811\text{-}821}$), yellow. b, Structure of a hibernating RNAP I dimer (PDB ID 4C2M). A135 subunit, black; A190 subunit, white; small subunits, beige. Inset, close-up view on the dimer interface. A190 clamp, violet; C-terminal A135 clamp, green.

4.15 Multi-angle light scattering and Negative stain EM analysis to confirm Dimeric (RNAP-δ-HelD)$_2$

Approximately two thirds of our cryoEM particle images represent RNAP-δ-HelD dimers. *In vivo* RNAP dimerization of eukaryotic RNA Pol I has been found (Torreira, E., et al., 2017). In addition, the dimerization of bacterial ribosomes has also been found (Beckert, B. et al., 2017; Khusainov, I. et al., 2017). This dimerization occurs when the cell is under stress conditions, and causes the eukaryotic RNA Pol I or bacterial ribosomes to enter the hibernating state. Thus, the dimeric Bs (RNAP--δ-HelD)$_2$ may be a hibernating state of Bs RNAP. In order to analyze if the dimerization we see in the cryoEM images is caused by the addition of detergent, I performed size exclusion chromatography runs combined with multi-angle light scattering (SEC-MALS) with Dr. Yongheng Huang and negative staining EM analysis of RNAP-δ-HelD in presence or absence of 0.15 % n-octylglucoside with Dr. Tarek Hilal.



M [kDa]	- OG	+ OG	Theor.
Peak 1 / Mon.	415	433	454.7
Peak 2 / Dimer	-	836	909.4

Figure 4.16: Investigation of RNAP-δ-HelD dimerization. a, SEC/multi-angle light scattering analysis of RNAP-δ-HelD used for cryoEM analysis in buffer lacking (-OG) or containing (+OG) 0.15 % (w/v) n-octylglucoside (critical micellar concentration 0.6 % [w/v]). Black traces, UV signals; red lines, molecular mass estimates across the peaks. Molecular masses deduced are listed in the bottom table compared to the theoretical (theor.) molecular masses for RNAP-δ-HelD (Mon.) and (RNAP-δ-HelD)2 (Dimer). About 16 % of the sample traverses the column as intact (RNAP-δ-HelD)2 dimers in the presence of n-octylglucoside. b, Top: negative stain EM micrographs of RNAP-δ-HelD in buffer lacking (-OG) or containing (+OG) 0.15 % (w/v) n-octylglucoside. Scale bars, 100 nm. Bottom: 2D class averages of picked particle images. Classes boxed red unequivocally indicate the presence of $(RNAP\text{-}\delta\text{-}HelD)_2$ dimers in both samples (with Huang, Y.H. and Hilal, T.).

Two peaks could be observed In the SEC-MALS analysis that correspond to the dimeric and monomeric conformation of the complex, respectively (Figure 4.16 a). Peak1, with a calculated molecular mass of ~433 kDa, represents the monomer's peak. Peak2, with a calculated molecular mass of ~836 kDa, represents the dimer's peak. The $(RNAP\text{-}\delta\text{-}HelD)_2$ dimers were partially stable during SEC under conditions

identical to cryoEM sample preparation (0.15 % n-octylglucoside; Figure 4.16 a). About 16% are dimeric RNAP-δ-HelD, while 84% are monomeric.

In negative stain EM analyses, we detected dimers both in the presence or absence of 0.15 % n-octylglucoside (Figure 4.16 b; a quantitative analysis of the monomer/dimer distribution was precluded by preferred particle orientations on the carbon films).

5 Discussion

Transcription is a fundamental process in gene expression, that is important in all kingdoms of life. Transcription is carried out by RNA polymerase, which is precisely regulated by relevant factors, such as σ in initiation, NusA and NusG in elongation. However, RNAP can be stuck on DNA templates or RNAs after termination. To reinitiate another round of transcription, RNAP must be recycled from the DNA template or the RNA.

5.1 RNAP recycling factor, HelD

Our results show that *B. subtilis* HelD has a two-pronged insert at the RNAP main and secondary channels. HelD clamps RNAP around the BH, widens the main channel and the RNA exit channel to provide escape routes for DNA and RNA, thereby releasing the bound nucleic acids. HelD uses arms of similar size to penetrate deeply into the channels, and δ plays a supporting role. δ^{NTD} helps to open the main channel, and δ^{CTR} may support to recruit HelD and guide $HelD^{Bumper}$ into the main channel to avoid topological trapping of DNA. Since HelD and δ can displace nucleic acids from RNAP without adding ATP, we speculate that the large interface area is buried upon RNAP-δ-HelD complex formation instead of HelD ATPase, which offers the dramatic driving force of RNAP opening.

In order to bind to RNAP, HelD reaches around the β2 lobe. This attack mode is not possible for RNAPs containing a β' lineage-specific insertion, SI3, stacked onto the β2 lobe, such as *E. coli* (Figure 5.1 a). Consistently, *E. coli* does not have HelD, but a far related ATPase, RapA, has been proposed to assist RNAP recycling (Sukhodolets, M.V. et al., 2001). Unlike HelD, RapA binds near the RNA exit channel and does not induce dramatic conformational changes in the EC (Figure 5.1 b). On the contrary, RapA is believed to rescue ECs via promoting backtracking (Liu, B. et al., 2015). There may be other recycling mechanisms in SI3-containing species. In fact, *E. coli* DksA has recently been found to remove RNAP from bound nucleic acids (Myka, K.K. and Gottesman, M.E. 2019). DksA uses a coiled-coil similar to GreA/B factors to bind in the secondary channel of RNAP (Perederina, A. et al., 2004) and induces RNAP conformational changes (Molodtsov, V. et al., 2018). The changes are not as dramatical as the effects of HelD and they are only found in bacteria containing SI3 in their RNAP (Furman, R. et al., 2013).

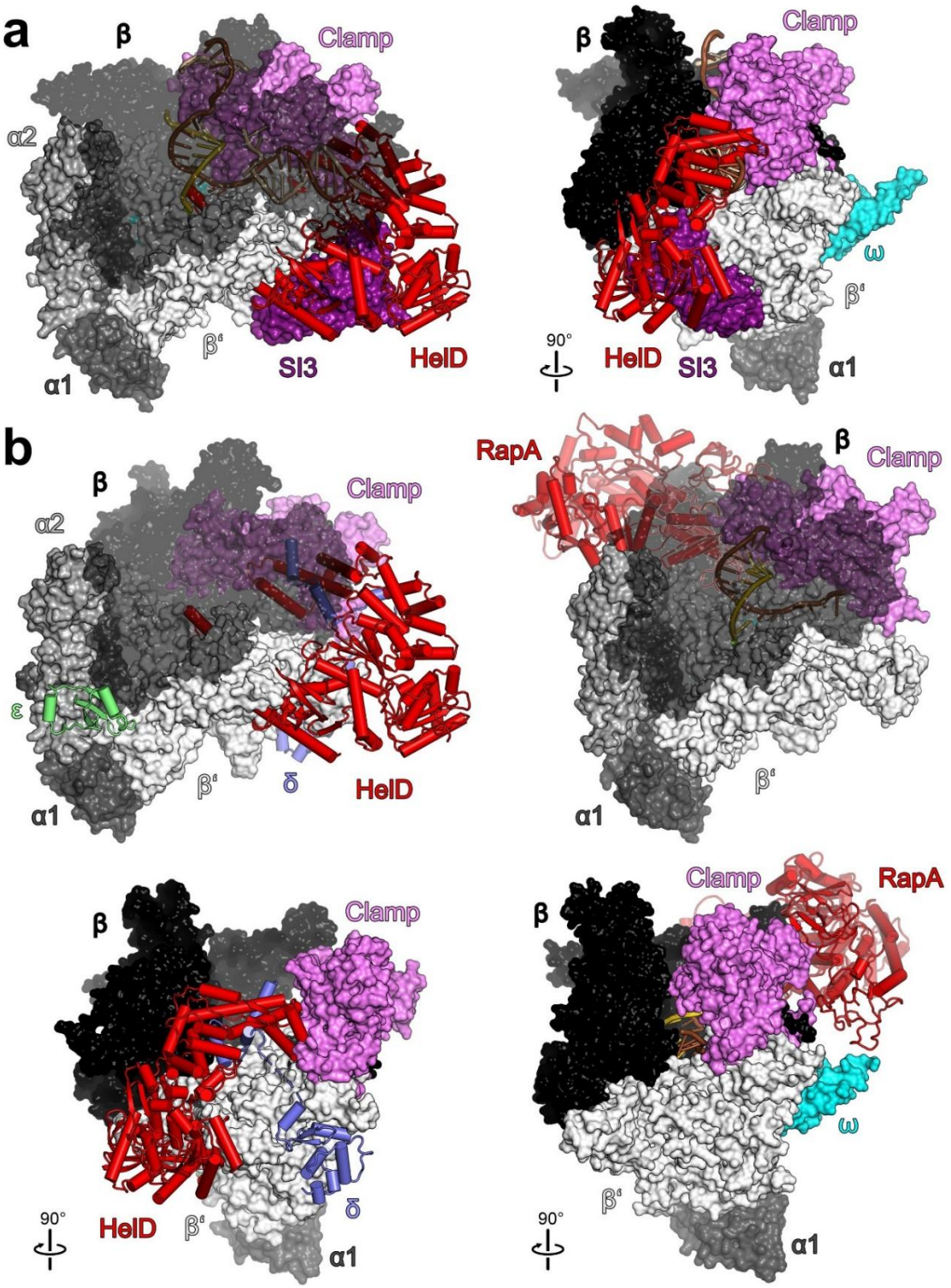


Figure 5.1 Comparison to SI3 RNAPs and a RapA complex. a, Orthogonal views of HelD transferred on an *E. coli* EC (PDB ID 6ALH) by superpositioning of the β subunits, showing that HelD would clash with SI3. SI3, purple. b, Comparison of the RNAP-δ-HelD complex (left) to an *E. coli* RapA-modified EC (right; PDB ID 4S20), illustrating how HelD and RapA attack the enzyme from opposite flanks (top), and that RapA binds RNAP in an EC-like closed conformation (bottom).

5.2 Function of δ

Our cryoEM structures also provide information about the possible mechanisms of action of the δ subunit during transcription initiation and elongation. In previous studies, δ alone had been shown to displace nucleic acids from RNAP (López de Saro, F.J. et al., 1995), what we recapitulated here as well (Figure 4.12). Since δ^{CTR} peptides showed similar activity when they were added in excess and δ^{NTD} was found to bind RNAP, it was suggested that δ-mediated nucleic acid displacement involves δ^{NTD}-dependent tethering of the polyanionic δ^{CTR} to the core of RNAP (López de Saro, F.J. et al., 1995). Our cryoEM structures confirm and further refine this hypothesis. δ^{NTD} anchors δ^{CTR} on the edge of the main channel; because of its length and intrinsic disorder, δ^{CTR} can enter into and explore most of the space in the main channel, competing spatially and electrostatically with bound nucleic acids. This mode of action would explain how δ enhances the recycling of core RNAP in multi-round transcription assays (López de Saro, F.J. et al., 1995; Juang, Y.L. and Helmann, J.D. 1994), and it may constitute the major recycling mechanism in bacteria that contains δ but without HelD. Our data also provide an explanation for the finding that in the presence of σ factors, RNAP-δ still binds to the promoters and forms into closed complexes, but cannot establish contacts with the downstream DNA (Juang, Y.L. and Helmann, J.D. 1994; Achberger, E.C. et al., 1982; Chen, Y.F. and Helmann, J.D. 1997), which is required for transition to an open complex. Finally, the model indicates that elimination of the positively charged region at the δ^{CTR} N-terminus promotes more extended conformations of the CTR (Kubáň, V. et al., 2019), effectively eliminating the ability to inhibit δ^{CTR} from invading the main channel and thus reconciling increased effects of the CTR variants at promoters that form unstable complexes (Kubáň, V. et al., 2019).

δ and σ^{A} show a negative cooperativity, and δ tends to exchange it for other σ factors lacking σ1.1 (López de Saro, F.J. et al., 1999; Hyde, E.I. et al., 1986). In the *E. coli* σ^{70} holoenzyme, σ1.1 can reside in the main channel to prevent double- or single-stranded DNA from entering into the active site of RNAP (Chen, J. et al., 2019) (Figure 4.9 b middle). In order to allow DNA loading, the clamp must be further opened (Chakraborty, A. et al., 2012) or σ1.1 has to be moved (Mekler, V. et al., 2002). These observations indicate that the binding competition between δ and σ1.1 is fully in line with our structures (Figure 4.9 b). However, although δ^{NTD} is similar to the globular domain of σ1.1 (Zachrdla, M. et al., 2017), our results show that the structure-independent CTR (together with $HelD^{Bumper}$ if present) constitutes the σ1.1-competitive element that can occupy an equivalent area in the main channel of RNAP (Figure 4.9 b left).

5.3 HelD-δ/ω competition

Interestingly, in our RNAP-δ-HelD EM sample which RNAP supplemented with

recombinantly purified HelD and δ, the ω subunit was underrepresented (Figure 4.3 d). As mentioned earlier (Zachrdla, M. et al., 2017), the first ~70 residues of δ^{NTD} fold very similar to the globular domain of σ1.1 regions of group 1 σ factors (Schwartz, E.C. et al., 2008). However, unlike the σ1.1 domain in an *E. coli* σ^{70} holoenzyme (Chen, J. et al., 2019), δ^{NTD} does not enter into the main channel but sits on the surface of RNAP between the β' shelf and jaw (Figure 4.7; Figure 5.2). This is consistent with previously published *in vivo* crosslinking/mass spectrometry (CLMS) results (de Jong, L. et al., 2017). Compared to the *E. coli* EC (Kang, J.Y. et al., 2017), δ^{NTD} seems to help in opening the main channel by somewhat shrinking the jaw and β' shelf. Moreover, the opening of RNAP and slight displacement of the shelf mediated by δ^{NTD} leads to repositioning of β' secondary channel elements, which would conflict with ω at its canonical binding site (Figure 5.2). This explains the loss of ω in RNAP-δ and RNAP-δ-HelD complexes (Figure 4.1 e; Figure 4.3 d). The lack of continuous cryoEM density exceeding δ^{NTD} indicates that δ^{CTR} is suspended from the edge of the main channel in a flexible manner.

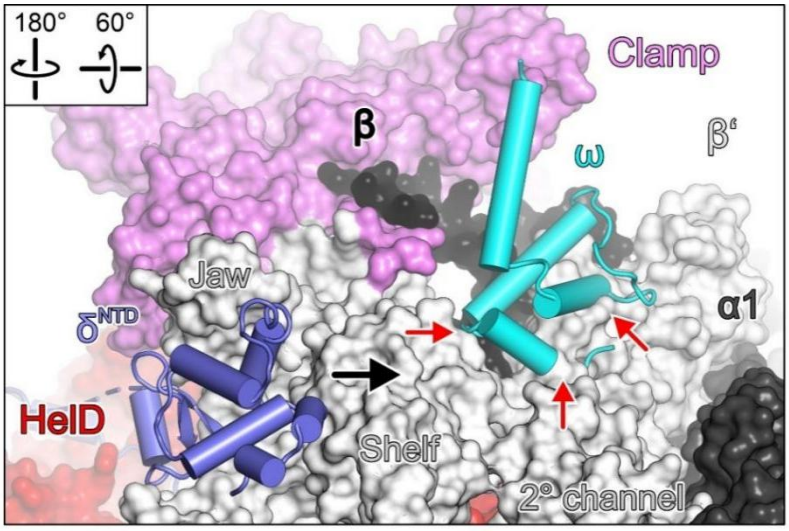


Figure 5.2 Closeup view on δ^{NTD} and ω binding regions with ω transferred from an *E. coli* EC (PDB ID 6ALH) by superpositioning of the β subunits. δ^{NTD} and HelD displace the β' shelf and other secondary channel elements (black arrow), leading to steric hindrance of ω binding (red arrows).

5.4 ε subunit

The position of ε subunit in *B. subtilis* RNAP is similar to a small domain in archaeal and eukaryotic nuclear RNAPs, which are homologs of the bacterial α1 subunit (D, Rpb3 and AC40 of archaeal RNAP, eukaryotic RNAP II and eukaryotic RNAP I/III, respectively; Figure 5.3). In some archaeal and eukaryotic RNAPs, these small domains bind an 4Fe-4S cluster (Hirata, A. et al., 2008), dashed circle (Figure 5.3), cavity in *B. subtilis* RNAP. However, this is not the case in the *E. coli* RNAP, which could in theory accommodate the equivalent archaeal subunit N (Rpb10 in eukaryotic

RNAP I, II and III), but it is not visible in the currently available structures. ε might support the structural integrity of RNAP. When HelD forces β and β' apart, it can ensure the interactions between α, β and β' subunits.

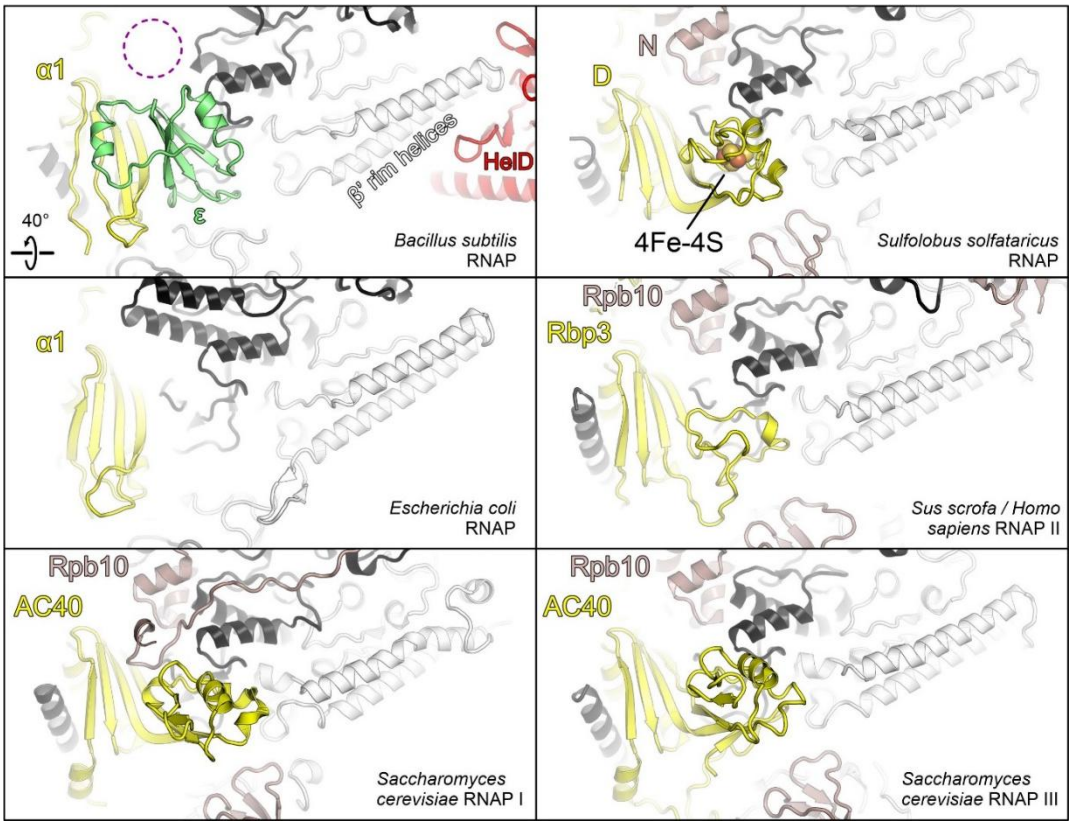


Figure 5.3 Comparison of ε to homologs. α1, α2, β, β' subunits and their homologs are colored yellow, gray, black and white, respectively; ε subunit, lime; HelD, red; subunits specific to archaeal and eukaryotic RNAPs, dark salmon. The illustration was prepared using structures with PDB IDs 3HKZ, 6ALH, 6GMH, 4C2M and 6TUT.

5.5 ATP dependent HelD release

The here unveiled HelD/δ-dependent recycling mechanism unveiled here shows an extremely simple, direct and effective way of resuming RNAP from almost any state of trapped post-termination. However, only when HelD also disassociates, the RNAP is truly recycled. We found that HelD is released by ATP (Figure 4.14 c), indicating that high levels of ATP could help to prevent HelD from trapping RNAP, which leads RNAP to an inactivated complex in exponential growth. It is worth noting that ATP cannot be bound when HelD is fully engaged in its interaction with RNAP, indicating that intrinsically timed isomerization into a less engaged conformation must precede ATP binding and release from RNAP. Unlike HelD, δ cannot bind ATP, and therefore ATP cannot be expected to directly modulate δ's affinity for RNAP. Consistently, ATP does not induce the accompanying release of δ (Figure 4.14 a), suggesting that δ has intrinsically high affinity to RNAP and does not require HelD to stably associate with

RNAP. As association of alternative σ factors (relative to σ^A) is more favored in RNAP-δ than in RNAP lacking δ (López de Saro, F.J. et al., 1999; Chakraborty, A. et al., 2012), additional steps might play a role to remove δ (or displace δ^{CTR} from the main channel) in the case of a special need to effectively bind σ^A. In any case, we suggest that ATP-mediated HelD release is the basis of the ATP-dependent stimulatory effect of HelD on multiple-round transcription (Wiedermannová, J. et al., 2014).

5.6 Dimerization and Hibernation?

Conversely, when *B. subtilis* cells sporulate during stationary phase, the levels of ATP are low (Hutchison, K.W. and Hanson, R.S. 1974), transcription is restricted, HelD levels match levels of RNAP (Nicolas, P. et al., 2012), and HelD is therefore expected to keep binding to RNAP. Does HelD keeping RNAP in an inactive state provide a mechanism to store RNAP until the environmental living conditions have improved? Interestingly, we found $(RNAP\text{-}\delta\text{-}HelD)_2$ dimers look very similar to the hibernating eukaryotic RNAP I (Figure 4.15). These dimers were partially stable in SEC at initial RNAP concentrations about 10-fold lower than their nominal cellular concentrations in the log phase, estimated from transcript levels and ribosome profiling (Nicolas, P. et al., 2012; Lalanne, J.B. et al., 2018). Furthermore, dimerization of RNAP has also been found in bacteria that lack HelD, including *E. coli* (Busby, S. et al., 1981). Therefore, dimerization might be an intrinsic property of RNAPs. In the presence of HelD, we observed a RNAP dimerization mode that involves homologous interactions between the β' clamps, the C-terminal β clamp and regions of the β flap; clearly, the HelD pushing the β' clamp outwards leads to the concomitant usage of these contact points (Figure 4.15 a). Thus, although not directly participating in forming the interface of the dimer, HelD clearly stimulates the observed dimerization. It is worth noting that comparing our dimeric structure to a *M. smegmatis* RNAP-σ^A holoenzyme structure (Kouba, T. et al., 2019) shows that all σ binding sites, except for σ1.1 in the main channel, could be accessible in the RNAP-δ-HelD dimer. Therefore, re-binding of σ might contribute to effective recovery of RNAP from the dimeric state. Taken together, HelD/δ could in principle promote RNAP dormancy, which might be necessary for rapid RNAP recovery. This is consistent with findings that overexpression of HelD enhances sporulation (Meeske, A.J. et al., 2016) and deletions of HelD, δ or both extend the lag phase (Wiedermannová, J. et al., 2014). To verify this idea, further tests are required. We could try *in vivo* CLMS at different growth phases and especially during sporulation in comparing WT to *ΔhelD* or overexpressing cells. Alternatively, we could try *in vivo* super-resolution imaging with fluorescence labeling of HelD or RNAP.

5.7 Model for HelD/δ-mediated RNAP recycling and putative hibernation.

Our research proposes a hitherto unrecognized transcription recycling system that protects genome integrity and might contribute to persistence during the hibernation state. In our model (Figure 5.4), of which parts require further validation, reservoirs of active RNAP are regulated by HelD, which may rescue trapped RNAP from DNA/RNA during rapid growth, promote RNAP dormancy during slow growth, and enable efficient RNAP recovery upon shift to a nutrient-rich environment. We note that although, for convenience, most laboratory experiments are performed with rapidly growing bacteria, hibernating states are common in natural environments and pose a serious health risk. For example, the spores of *B. anthracis* are the infectious particles for anthrax, whereas slow-growing *Pseudomonas aeruginosa* biofilms and *M. tuberculosis* are resistant to cidal antibiotics. Therefore, understanding the regulation of hibernation is critical for understanding bacterial physiology and identifying new strategies for eradication of multidrug-resistant pathogens.

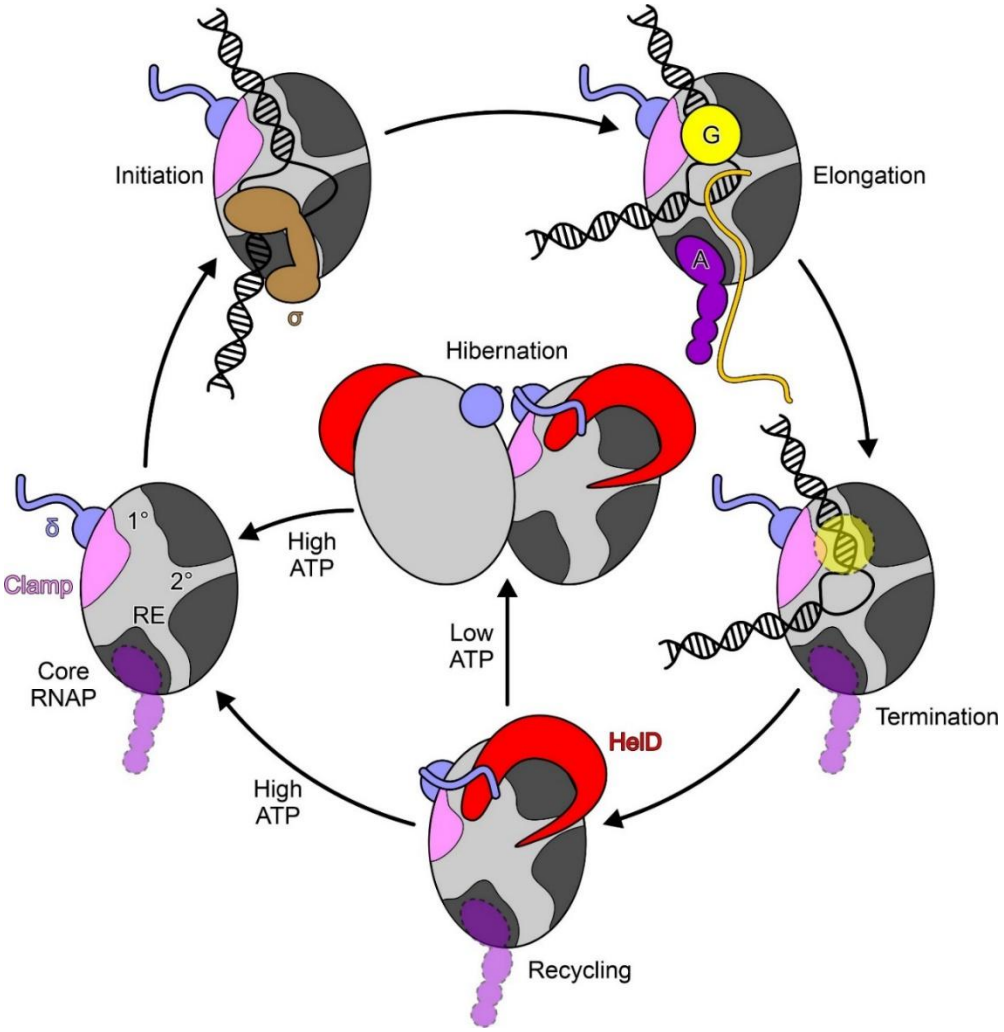


Figure 5.4 Model for HelD/δ-mediated RNAP recycling and putative hibernation. 1°/2°, main/secondary channels; RE, RNA exit tunnel; A/G, general elongation factors NusA/NusG. NusG binds across the active center cleft, while NusA binds to the β FT. Semitransparent icons with dashed lines indicate that the respective factor may be released at the respective step. If the factors remain after termination, NusG will likely be displaced by HelD-induced main channel opening, while the NusA binding site is sequestered in hibernating RNAP-HelD. Hibernation by RNAP-δ-HelD dimerization and ATP-mediated recovery from the dormant state represent tentative aspects of the model.

6 References

Achberger, E.C., Hilton, M.D., Whiteley, H.R. The effect of the delta subunit on the interaction of Bacillus subtilis RNA polymerase with bases in a SP82 early gene promoter. *Nucleic Acids Res* **10**, 2893-910 (1982).

Archambault, J. & Friesen, J.D. Genetics of RNA polymerases I, II, and III. *Microbiol Rev* **57**, 703–724 (1993).

Altmann, C.R., Solow-Cordero, D.E., Chamberlin, M.J. RNA cleavage and chain elongation by Escherichia coli DNA-dependent RNA polymerase in a binary enzyme.RNA complex. *Proc Natl Acad Sci U S A* **91**, 3784-8 (1994).

Artsimovitch, I. & Landick, R. The transcriptional regulator RfaH stimulates RNA chain synthesis after recruitment to elongation complexes by the exposed nontemplate DNA strand. *Cell* **109**, 193-203 (2002).

Afonine, P.V., Grosse-Kunstleve, R.W., Echols, N., Headd, J.J., Moriarty, N.W., Mustyakimov, M., Terwilliger, T.C., Urzhumtsev, A., Zwart, P.H., Adams, P.D. Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr D* **68**, 352-367 (2012).

Abdelkareem, M., Saint-André, C., Takacs, M., Papai, G., Crucifix, C., Guo, X., Ortiz, J., Weixlbaumer, A. Structural Basis of Transcription: RNA Polymerase Backtracking and Its Reactivation. *Mol Cell* **75**, 298-309 e4 (2019).

Bernecky, C., Herzog, F., Baumeister, W., Plitzko, J.M., Cramer, P. Structure of transcribing mammalian RNA polymerase II. *Nature* **529**, 551-554 (2016).

Busby, S., Spassky, A., Buc, H. On the binding of tRNA to Escherichia coli RNA polymerase. Interactions between the core enzyme, DNA and tRNA. *Eur J Biochem* **118**, 443-51 (1981).

Boylan, S. A., Redfield, A.R., Brody, M.S., Price, C.W. Stress induced activation of the sigma B transcription factor of Bacillus subtilis. *J. Bacteriol.* **175**, 7931–7937 (1993).

Borukhov, S., Sagitov, V., Goldfarb, A. Transcript cleavage factors from E. coli. *Cell* **72**, 459-66 (1993).

Borukhov, S. and Nudler, E. RNA polymerase: the vehicle of transcription. *Trends Microbiol* **16,** 126-134 (2008).

Beckert, B., Abdelshahid, M., Schäfer, H., Steinchen, W., Arenz, S., Berninghausen, O., Beckmann, R., Bange, G., Turgay, K., Wilson, D.N. Structure of the *Bacillus subtilis* hibernating 100S ribosome reveals the basis for 70S dimerization. *EMBO J* **36**, 2061-2072 (2017).

Bellecourt, M.J., Ray-Soni, A., Harwig, A., Mooney, R.A., Landick, R. RNA Polymerase Clamp Movement Aids Dissociation from DNA but Is Not Required for RNA Release at Intrinsic Terminators. *J Mol Biol* **431**, 696-713 (2019).

Belogurov, G.A. & Artsimovitch, I. The Mechanisms of Substrate Selection, Catalysis, and Translocation by the Elongating RNA Polymerase. *J Mol Biol* **431**, 3975-4006 (2019).

Burmann, B.M., Schweimer, K., Luo, X., Wahl, M.C., Stitt, B.L., Gottesman, M.E., Rösch, P. A NusE:NusG complex links transcription and translation. *Science* **328**, 501–504 (2010).

Chang, B. Y. & Doi, R.H. Overproduction, purification, and characterization of Bacillus subtilis RNA polymerase sigma A factor. *J. Bacteriol.* **172**, 3257–3263 (1990).

Chen, Y.F. & Helmann, J.D. DNA-melting at the Bacillus subtilis flagellin promoter nucleates near -10 and expands unidirectionally. *J Mol Biol* **267**, 47-59 (1997).

Ciampi, M.S. Rho-dependent terminators and transcription termination. *Microbiology* **152**, 2515-2528 (2006).

Cramer, P., Bushnell, D.A., Kornberg, R.D. Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. *Science* **292,** 1863-1876 (2001).

Chen, Z.A., Jawhari, A., Fischer, L., Buchen, C., Tahir, S., Kamenski, T., Rasmussen, M., Lariviere, L., Bukowski-Wills, J.C., Nilges, M., Cramer, P., Rappsilber, J. Architecture of the RNA polymerase II-TFIIF complex revealed by cross-linking and mass spectrometry. *EMBO J* **29**, 717-26 (2010).

Chakraborty, A., Wang, D., Ebright, Y.W., Korlann, Y., Kortkhonjia, E., Kim, T., Chowdhury, S., Wigneshweraraj, S., Irschik, H., Jansen, R., Nixon, B.T., Knight, J., Weiss, S., Ebright, R.H. Opening and closing of the bacterial RNA polymerase clamp. *Science* **337**, 591-5 (2012).

Chen, J., Wassarman, K.M., Feng, S., Leon, K., Feklistov, A., Winkelman, J.T., Li, Z., Walz, T., Campbell, E.A., Darst, S.A. 6S RNA Mimics B-Form DNA to Regulate Escherichia coli RNA Polymerase. *Mol Cell* **68**, 388-397 (2017).

Chen, J., Gopalkrishnan, S., Chiu, C., Chen, A.Y., Campbell, E.A., Gourse, R.L., Ross, W., Darst, S.A. E. coli TraR allosterically regulates transcription initiation by altering RNA polymerase conformation. *Elife* **8** (2019).

d'Aubenton Carafa, Y., Brody, E., Thermes, C. Prediction of rho-independent Escherichia coli transcription terminators. A statistical analysis of their RNA stem-loop structures. *J Mol Biol* **216**, 835-858 (1990).

Deaconescu, A.M., Chambers, A.L., Smith, A.J., Nickels, B.E., Hochschild, A., Savery, N.J., Darst, S.A. Structural basis for bacterial transcription-coupled DNA repair. *Cell* **124**, 507-20 (2006).

Delumeau, O., Lecointe, F., Muntel, J., Guillot, A., Guédon, E., Monnet, V., Hecker, M., Becher, D., Polard, P., Noirot, P. The dynamic protein partnership of RNA polymerase in Bacillus subtilis. *Proteomics* **11**, 2992-3001 (2011).

de Hoon, M.J.L., Makita, Y., Nakai, k., Miyano, S. Prediction of Transcriptional Terminators in Bacillus subtilis and Related Species. *PLoS Comp Biol* **1(3)**, e25 (2005).

de Jong, L., de Koning, E.A., Roseboom, W., Buncherd, H., Wanner, M.J., Dapic, I., Jansen, P.J., van Maarseveen, J.H., Corthals, G.L., Lewis, P.J., Hamoen, L.W., de Koster, C.G. In-Culture Cross-Linking of Bacterial Cells Reveals Large-Scale

Dynamic Protein-Protein Interactions at the Peptide Level. *J Proteome Res* **16**, 2457-2471 (2017).

Ederth, J., Mooney, R.A., Isaksson, L.A., Landick, R. Functional interplay between the jaw domain of bacterial RNA polymerase and allele-specific residues in the product RNA-binding pocket. *J Mol Biol* **356**, 1163-1179 (2006).

Epshtein, V., Cardinale, C.J., Ruckenstein, A.E., Borukhov, S., Nudler, E. An allosteric path to transcription termination. Molecular Cell 28, 991-1001 (2007).

Epshtein, V., Dutta, D., Wade, J., Nudler, E. An allosteric mechanism of Rho-dependent transcription termination. *Nature* **463**, 245–249 (2010).

Emsley, P., Lohkamp, B., Scott, W.G., Cowtan, K. Features and development of Coot. *Acta Crystallogr D* **66**, 486-501 (2010).

Engel, C., Sainsbury, S., Cheung, A.C., Kostrewa, D., Cramer, P. RNA polymerase I structure and transcription regulation. *Nature* **502**, 650-5 (2013).

Friedman, D.I. & Baron, L.S. Genetic characterization of a bacterial locus involved in the activity of the N function of phage lambda. *Virology* **58**, 141-8 (1974).

Ferrari, R., Rivetti, C., Dieci, G. Transcription reinitiation properties of bacteriophage T7 RNA polymerase. Biochem Biophys Res Commun 315, 376-80 (2004).

Fernández-Tornero, C., Moreno-Morcillo, M., Rashid, U.J., Taylor, N.M., Ruiz, F.M., Gruene, T., Legrand, P., Steuerwald, U., Müller, C.W. Crystal structure of the 14-subunit RNA polymerase I. *Nature* **502**, 644-9 (2013).

Furman, R., Tsodikov, O.V., Wolf, Y.I., Artsimovitch, I. An insertion in the catalytic trigger loop gates the secondary channel of RNA polymerase. *J Mol Biol* **425**, 82-93 (2013).

Fischer, L. & Rappsilber, J. Quirks of Error Estimation in Cross-Linking/Mass Spectrometry. *Anal Chem* **89**, 3829-3833 (2017).

Fernández-Tornero, C. RNA polymerase I activation and hibernation: unique mechanisms for unique genes. *Transcription* **9**, 248-254 (2018).

Gill, S.C., Weitzel, S.E., von Hippel, P.H. Escherichia coli sigma 70 and NusA proteins. I. Binding interactions with core RNA polymerase in solution and within the transcription complex. *J Mol Biol* **220,** 307-324 (1991).

Gusarov, I. & Nudler, E. The mechanism of intrinsic transcription termination. *Mol Cell* **3**, 495-504 (1999).

Gohara, D.W. & Yap, M.F. Survival of the drowsiest: the hibernating 100S ribosome in bacterial stress management. *Curr Genet* **64**, 753-760 (2018).

Hanawalt, P.C. & Haynes, R.H. The repair of DNA. *Sci Am.* **216**, 36-43 (1967).

Hutchison, K.W. & Hanson, R.S. Adenine nucleotide changes associated with the initiation of sporulation in Bacillus subtilis. *J Bacteriol* **119**, 70-5 (1974).

Hickson, I.D., Arthur, H.M., Bramhill, D., Emmerson, P.T. The E. coli uvrD gene product is DNA helicase II. *Mol Gen Genet.* **190**, 265-70 (1983).

Hyde, E.I., Hilton, M.D., Whiteley, H.R. Interactions of Bacillus subtilis RNA polymerase with subunits determining the specificity of initiation. Sigma and delta peptides can bind simultaneously to core. *J Biol Chem* **261**, 16565-70 (1986).

Hirata, A., Klein, B.J., Murakami, K.S. The X-ray crystal structure of RNA polymerase from Archaea. *Nature* **451**, 851-4 (2008).

Hubin, E.A., Lilic, M., Darst, S.A., Campbell, E.A. Structural insights into the mycobacteria transcription initiation complex from analysis of X-ray crystal structures. *Nat Commun* **8**, 16072 (2017).

Herbert, K.M., Zhou, J., Mooney, R.A., La Porta, A., Landick, R., Block, S.M. E. coli NusG inhibits backtracking and accelerates pause-free transcription by promoting forward translocation of RNA polymerase. *J Mol Biol* **399**, 17-30 (2010).

Helmann, J.D. Where to begin? Sigma factors and the selectivity of transcription initiation in bacteria. *Mol Microbiol* **112**, 335-347 (2019).

Harden, T.T., Herlambang K.S., Chamberlain M., Lalanne J.B., Wells C.D., Li G.W., Landick R., Hochschild A., Kondev J., Gelles J. Alternative transcription cycle for bacterial RNA polymerase. *Nat Commun* **11**, 448 (2020).

Ingham, C.J., Dennis, J., Furneaux, P.A. Autogenous regulation of transcription termination factor Rho and the requirement for Nus factors in Bacillus subtilis. *Mol Microbiol* **31** 651–663 (1999).

Ito, K., Iwakura, Y., Ishihama, A. Biosynthesis of RNA polymerase in Escherichia coli. III. Identification of intermediates in the assembly of RNA polymerase. *J Mol Biol* **96**, 257-271 (1975).

Juang, Y.L. & Helmann, J.D. The delta subunit of Bacillus subtilis RNA polymerase. An allosteric effector of the initiation and core-recycling phases of transcription. *J Mol Biol* **239**, 1-14 (1994).

Jeon, Y.H., Yamazaki, T., Otomo, T., Ishihama, A., Kyogoku, Y. Flexible linker in the RNA polymerase alpha subunit facilitates the independent motion of the C-terminal activator contact domain. *J Mol Biol* **267**, 953-962 (1997).

Jia, H., Korolev, S., Niedziela-Majka, A., Maluf, N.K., Gauss, G.H., Myong, S., Ha, T., Waksman, G., Lohman, T.M. Rotations of the 2B sub-domain of E. coli UvrD helicase/translocase coupled to nucleotide and DNA binding. *J Mol Biol* **411**, 633-48 (2011).

Johnson, G.E., Lalanne, J.B., Peters, M.L., Li, G.W. Functionally uncoupled transcription–translation in Bacillus subtilis. *Nature* **585**, 124-128 (2020).

King, R.A., Markov, D., Sen, R., Severinov, K., Weisberg, R.A. A conserved zinc binding domain in the largest subunit of DNA-dependent RNA polymerase modulates intrinsic transcription termination and antitermination but does not stabilize the elongation complex. *J Mol Biol* **342**, 1143-1154 (2004).

Kettenberger, H., Eisenführ, A., Brueckner, F., Theis, M., Famulok, M., Cramer, P. Structure of an RNA polymerase II-RNA inhibitor complex elucidates transcription regulation by noncoding RNAs. *Nat Struct Mol Biol* **13**, 44-8 (2006).

Kuhn, C.D., Geiger, S.R., Baumli, S., Gartmann, M., Gerber, J., Jennebach, S., Mielke, T., Tschochner, H., Beckmann, R., Cramer, P. Functional architecture of RNA polymerase I. *Cell* **131**, 1260-72 (2007).

Keller, A.N., Yang, X., Wiedermannová, J., Delumeau, O., Krásný, L., Lewis, P.J. ε, a new subunit of RNA polymerase found in gram-positive bacteria. *J Bacteriol* **196**, 3622-32 (2014).

Kakar, S., Fang, X., Lubkowska, L., Zhou, Y.N., Shaw, G.X., Wang, Y.X., Jin, D.J., Kashlev, M., Ji, X. Allosteric Activation of Bacterial Swi2/Snf2 (Switch/Sucrose Non-fermentable) Protein RapA by RNA Polymerase: BIOCHEMICAL AND STRUCTURAL STUDIES. *J Biol Chem* **290**, 23656-69 (2015).

Kang, J.Y., Olinares, P.D., Chen, J., Campbell, E.A., Mustaev, A., Chait, B.T., Gottesman, M.E., Darst, S.A. Structural basis of transcription arrest by coliphage HK022 Nun in an Escherichia coli RNA polymerase elongation complex. *Elife* **6**, e25478 (2017).

Khusainov, I., Vicens, Q., Bochler, A., Grosse, F., Myasnikov, A., Ménétret, J.F., Chicher, J., Marzi, S., Romby, P., Yusupova, G., Yusupov, M., Hashem, Y. Structure of the 70S ribosome from human pathogen Staphylococcus aureus. *Nucleic Acids Res* **45**, 1026 (2017).

Kovaľ, T., Sudzinová, P., Perháčová, T., Trundová, M., Skálová, T., Fejfarová, K., Šanderová, H., Krásný, L., Dušková, J., Dohnálek, J. Domain structure of HelD, an interaction partner of Bacillus subtilis RNA polymerase. *FEBS Lett* **593**, 996-1005 (2019).

Kouba, T., Pospíšil, J., Hnilicová, J., Šanderová, H., Barvík, I., Krásný, L. The Core and Holoenzyme Forms of RNA Polymerase from Mycobacterium smegmatis. *J Bacteriol* **201**(2019).

Kubáň, V., Srb, P., Štégnerová, H., Padrta, P., Zachrdla, M., Jaseňáková, Z., Šanderová, H., Vítovská, D., Krásný, L., Koval', T., Dohnálek, J., Ziemska-Legiecka, J., Grynberg, M., Jarnot, P., Gruca, A., Jensen, M.R., Blackledge, M., Žídek, L. Quantitative Conformational Analysis of Functionally Important Electrostatic Interactions in the Intrinsically Disordered Region of Delta Subunit of Bacterial RNA Polymerase. *J Am Chem Soc* **141**, 16817-16828 (2019).

Kang, W., Ha K.S., Uhm H., Park K., Lee J.Y., Hohng S., Kang C. Transcription reinitiation by recycling RNA polymerase that diffuses on DNA after releasing terminated RNA. *Nat Commun* **11**, 450 (2020).

Laemmli, U.K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685 (1970).

Landick, R. & Yanofsky, C. Isolation and structural analysis of the Escherichia coli trp leader paused transcription complex. *J Mol Biol* **196**, 363-377 (1987).

López de Saro, F.J., Woody, A.Y., Helmann, J.D. Structural analysis of the Bacillus subtilis delta factor: a protein polyanion which displaces RNA from RNA polymerase. *J Mol Biol* **252**, 189-202 (1995).

López de Saro, F.J., Yoshikawa, N., Helmann, J.D. Expression, abundance, and RNA polymerase binding properties of the delta factor of Bacillus subtilis. *J Biol Chem* **274**, 15953-8 (1999).

Lee, J.Y. & Yang, W. UvrD helicase unwinds DNA one base pair at a time by a two-part power stroke. *Cell* **127**, 1349-60 (2006).

Leitner, A., Walzthoeni, T., Aebersold, R. Lysine-specific chemical cross-linking of protein complexes and identification of cross-linking sites using LC-MS/MS and the xQuest/xProphet software pipeline. *Nat Protoc* **9**, 120-37 (2014).

Liu, B., Zuo, Y., Steitz, T.A. Structural basis for transcription reactivation by RapA. *Proc Natl Acad Sci U S A* **112**, 2006-10 (2015).

Lalanne, J.B., Taggart, J.C., Guo, M.S., Herzel, L., Schieler, A., Li, G.W. Evolutionary Convergence of Pathway-Specific Enzyme Expression Stoichiometry. *Cell* **173**, 749-761 e38 (2018).

Matson, S.W. & George, J.W. DNA helicase II of Escherichia coli. Characterization of the single-stranded DNA-dependent NTPase and helicase activities. *J Biol Chem.* **262**, 2066-76 (1987).

Mellon, I. & Hanawalt, P.C. Induction of the Escherichia coli lactose operon selectively increases repair of its transcribed DNA strand. *Nature* **342**, 95-8 (1989).

Miller Jr., O.L., Hamkalo, B.A., Thomas Jr., C.A. Visualization of bacterial genes in action. *Science* **169**, 392–395 (1970).

Milkereit, P., Schultz, P., Tschochner, H. Resolution of RNA polymerase I into dimers and monomers and their function in transcription. *Biol Chem* **378**, 1433-43 (1997).

Mukherjee, K., Nagai, H., Shimamoto, N., Chatterji, D. GroEL is involved in activation of Escherichia coli RNA polymerase devoid of the omega subunit in vivo. *Eur J Biochem* **266,** 228-235 (1999).

Mekler, V. Kortkhonjia, E., Mukhopadhyay, J., Knight, J., Revyakin, A., Kapanidis, A.N., Niu, W., Ebright, Y.W., Levy, R., Ebright, R.H. Structural organization of bacterial RNA polymerase holoenzyme and the RNA polymerase-promoter open complex. *Cell* **108**, 599-614 (2002).

Mooney, R.A., Davis, S.E., Peters, J.M., Rowland, J.L., Ansari, A.Z., Landick, R. Regulator trafficking on bacterial transcription units in vivo. *Mol Cell* **33,** 97-108 (2009).

Motácková, V., Sanderová, H., Zídek, L., Novácek, J., Padrta, P., Svenková, A., Korelusová, J., Jonák, J., Krásný, L., Sklenár, V. Solution structure of the N-terminal domain of Bacillus subtilis delta subunit of RNA polymerase and its classification based on structural homologs. *Proteins* **78**, 1807-10 (2010).

Meeske, A.J., Rodrigues, C.D., Brady, J., Lim, H.C., Bernhardt, T.G., Rudner, D.Z. High-Throughput Genetic Screens Identify a Large and Diverse Collection of New Sporulation Genes in Bacillus subtilis. *PLoS Biol* **14**, e1002341 (2016).

Molodtsov, V., Sineva, E., Zhang, L., Huang, X., Cashel, M., Ades, S.E., Murakami, K.S. Allosteric Effector ppGpp Potentiates the Inhibition of Transcript Initiation by DksA. *Mol Cell* **69**, 828-839 e5 (2018).

Myka, K.K. & Gottesman, M.E. DksA and DNA double-strand break repair. *Curr Genet* **65**, 1297-1300 (2019).

Mendes, M.L., Fischer, L., Chen, Z.A., Barbon, M., O'Reilly, F.J., Giese, S.H., Bohlke-Schneider, M., Belsom, A., Dau, T., Combe, C.W., Graham, M., Eisele, M.R., Baumeister, W., Speck, C., Rappsilber, J. An integrated workflow for crosslinking mass spectrometry. *Mol Syst Bio* **15**, e8994 (2019).

Nudler, E. & Gottesman, M.E. Transcription termination and anti-termination in E. coli. *Genes Cells* **7**, 755-768 (2002).

Nicolas, P., Noirot, P. et al. Condition-dependent transcriptome reveals high-level regulatory architecture in Bacillus subtilis. *Science* **335**, 1103–1106 (2012).

Neyer, S., Kunz, M., Geiss, C., Hantsche, M., Hodirnau, V.V., Seybert, A., Engel, C., Scheffer, M.P., Cramer, P., Frangakis, A.S. Structure of RNA polymerase I transcribing ribosomal DNA genes. *Nature* **540**, 607-610 (2016).

Petersohn, A., Brigulla, M., Haas, S., Hoheisel, J.D., Volker, U., Hecker, M. Global analysis of the general stress response of Bacillus subtilis. *J. Bacteriol.* **183**, 5617–5631 (2001).

Perederina, A., Svetlov, V., Vassylyeva, M.N., Tahirov, T.H., Yokoyama, S., Artsimovitch, I., Vassylyev, D.G. Regulation through the secondary channel--structural framework for ppGpp-DksA synergism during transcription. *Cell* **118**, 297-309 (2004).

Papoušková, V., Kadeřávek, P., Otrusinová, O., Rabatinová, A., ŠSanderová, H., Nováček, J., Krásný, L., Sklenář, V., Žídek, L. Structural study of the partially disordered full-length δ subunit of RNA polymerase from Bacillus subtilis. *Chembiochem* **14**, 1772-9 (2013).

Pilsl, M., Crucifix, C., Papai, G., Krupp, F., Steinbauer, R., Griesenbeck, J., Milkereit, P., Tschochner, H., Schultz, P. Structure of the initiation-competent RNA polymerase I and its implication for transcription. *Nat Commun* **7**, 12126 (2016).

Punjani, A., Rubinstein, J.L., Fleet, D.J. & Brubaker, M.A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat Methods* **14**, 290-296 (2017).

Qi, Y. & Hulett, F.M. PhoP-P and RNA polymerase sigmaA holoenzyme are sufficient for transcription of Pho regulon promoters in Bacillus subtilis: PhoP-P activator sites within the coding region stimulate transcription in vitro. *Mol Microbiol* **28**, 1187-97 (1998).

Ryan, T. & Chamberlin, M.J. Transcription analyses with heteroduplex trp attenuator templates indicate that the transcript stem and loop structure serves as the termination signal. *J Biol Chem* **258**, 4690-3 (1983).

Ramsay, E.P., Abascal-Palacios, G., Daiß, J.L., King, H., Gouge, J., Pilsl, M., Beuron, F., Morris, E., Gunkel, P., Engel, C., Vannini, A. Structure of human RNA polymerase III. *Nat Commun* **11**, 6409 (2020).

Rappsilber, J., Mann, M., Ishihama, Y. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat Protoc* **2**, 1896-906 (2007).

Rabatinová, A., Šanderová, H., Jirát Matějčková, J., Korelusová, J., Sojka, L., Barvík, I., Papoušková, V., Sklenár, V., Žídek, L., Krásný, L. The δ subunit of RNA

polymerase is required for rapid changes in gene expression and competitive fitness of the cell. *J Bacteriol* **195**, 2603-11 (2013).

Said, N., Hilal, T., Sunday, N.D., Khatri, A., Bürger, J., Mielke, T., Belogurov, G.A., Loll, B., Sen, R.J., Artsimovitch, I., Wahl, M.C. Steps toward translocation-independent RNA polymerase inactivation by terminator ATPase ρ. *Science*, 10.1126/science.abd1673 (2020).

Saxena, S., Myka, K.K., Washburn, R., Costantino, N., Court, D.L., Gottesman, M.E. Escherichia coli transcription factor NusG binds to 70S ribosomes. *Mol. Microbiol*. **108**, 495–504 (2018).

Shorenstein, R. G. & R. Losick. Purification and properties of the sigma subunit of ribonucleic acid polymerase from vegetative Bacillus subtilis. *J. Biol. Chem.* **248**, 6163–6169 (1973).

Sambrook, J. & Fritsch, E.F. Molecular cloning-A laboratory manual. *Cold Spring Harbour, New York, Cold Spring Harbour Laboratory Press* (1989).

Selby, C.P. & Sancar, A. Transcription preferentially inhibits nucleotide excision repair of the template DNA strand in vitro. *J Biol Chem* **265**, 21330-6 (1990).

Selby, C.P. & Sancar, A. Gene- and strand-specific repair in vitro: partial purification of a transcription-repair coupling factor. *Proc Natl Acad Sci U S A* **88**, 8232-6 (1991).

Selby, C.P. & Sancar, A. Molecular mechanism of transcription-repair coupling. *Science* **260**, 53-8 (1993).

Svoboda, D.L., Smith, C.A., Taylor, J.S., Sancar, A. Effect of sequence, adduct type, and opposing lesions on the binding and repair of ultraviolet photodamage by DNA photolyase and (A)BC excinuclease. *J Biol Chem* **268**, 10694-700 (1993).

Sukhodolets, M.V., Jin D.J. RapA, a novel RNA polymerase-associated protein, is a bacterial homolog of SWI2/SNF2. *J Biol Chem* **273**, 7018-7023 (1998).

Sukhodolets, M.V., Cabrera, J.E., Zhi, H., Jin, D.J. RapA, a bacterial homolog of SWI2/SNF2, stimulates RNA polymerase recycling in transcription. *Genes Dev* **15**, 3330-41 (2001).

Studier, F.W. Protein production by auto-induction in high density shaking cultures. *Protein Expr Purif* **41**, 207-34 (2005).

Shaw, G., Gan, J., Zhou, Y.N., Zhi, H., Subburaman, P., Zhang, R., Joachimiak, A., Jin, D.J., Ji, X. Structure of RapA, a Swi2/Snf2 protein that recycles RNA polymerase during transcription. *Structure* **16**, 1417-27 (2008).

Schwartz, E.C., Shekhtman, A., Dutta, K., Pratt, MR., Cowburn, D., Darst, S., Muir, T.W. A full-length group 1 bacterial sigma factor adopts a compact structure incompatible with DNA binding. *Chem Biol* **15**, 1091-103 (2008).

Santangelo, T.J. & Artsimovitch I. Termination and antitermination: RNA polymerase runs a stop sign. *Nat Rev Microbiol* **9**, 319-329 (2011).

Selby, C.P. Mfd Protein and Transcription-Repair Coupling in Escherichia coli. *Photochem Photobiol* **93**, 280-295 (2017).

Thomsen, N.D., Lawson, M.R., Witkowsky, L.B., Qu, S., Berger, J.M. Molecular mechanisms of substrate-controlled ring dynamics and substepping in a nucleic

acid-dependent hexameric motor. *Proc. Natl. Acad. Sci. U.S.A.* **113**, E7691–E7700 (2016).

Travers, A. RNA polymerase specificity and the control of growth. *Nature* **263**, 641-6 (1976).

Tjian, R., Losick, R., Pero, J., Hinnebush, A. Purification and comparative properties of the delta and sigma subunits of RNA polymerase from Bacillus subtilis. *Eur. J. Biochem.* **74**, 149–154 (1977).

Torreira, E., Louro, J.A., Pazos, I., González-Polo, N., Gil-Carton, D., Duran, A.G., Tosi, S., Gallego, O., Calvo, O., Fernández-Tornero, C. The dynamic assembly of distinct RNA polymerase I complexes modulates rDNA transcription. *Elife* **6**, e20832 (2017).

Vassylyev, D.G., Sekine, S., Laptenko, O., Lee, J., Vassylyeva, M.N., Borukhov, S., Yokoyama, S. Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 A resolution. *Nature* **417,** 712-719 (2002).

Vrentas, C.E., Gaal, T., Ross, W., Ebright, R.H., Gourse, R.L. Response of RNA polymerase to ppGpp: requirement for the omega subunit and relief of this requirement by DksA. *Genes Dev* **19**, 2378-2387 (2005).

Vassylyev, D.G., Vassylyeva, M.N., Zhang, J., Palangat, M., Artsimovitch, I., Landick, R. Structural basis for substrate loading in bacterial RNA polymerase. *Nature* **448,** 163-168 (2007).

Wang, C.Y., Molodtsov, V., Firlar, E., Kaelber, J.T., Blaha, G., Su, M., Ebright, R.H. Structural basis of transcription-translation coupling. *Science* 10.1126/science.abb5317 (2020).

Wang, D., Bushnell D.A., Westover K.D., Kaplan C.D., Kornberg R.D. Structural basis of transcription: role of the trigger loop in substrate specificity and catalysis. *Cell* **127**, 941-954 (2006).

Wassarman, K.M. & Saecker, R.M. Synthesis-mediated release of a small RNA inhibitor of RNA polymerase. *Science* **314**, 1601-3 (2006).

Washburn, R.S. & Gottesman, M.E. Transcription termination maintains chromosome integrity. *Proc Natl Acad Sci U S A* **108**, 792-7 (2011).

Webster, M.W., Takacs, M., Zhu, C.J., Vidmar, V., Eduljee, A., Abdelkareem, M., Weixlbaumer, A., Structural basis of transcription-translation coupling and collision in bacteria. *Science* 10.1126/science.abb5036 (2020).

Werner, F. A nexus for gene expression-molecular mechanisms of Spt5 and NusG in the three domains of life. *J Mol Biol* **417**, 13-27 (2012).

Wiedermannová, J., Sudzinová, P., Kovaľ, T., Rabatinová, A., Šanderova, H., Ramaniuk, O., Rittich, Š., Dohnálek, J., Fu, Z., Halada, P., Lewis, P., Krásny, L. Characterization of HelD, an interacting partner of RNA polymerase from Bacillus subtilis. *Nucleic Acids Res* **42**, 5151-63 (2014).

Williams, C.J., Headd, J.J., Moriarty, N.W., Prisant, M.G., Videau, L.L., Deis, L.N., Verma, V., Keedy, D.A., Hintze, B.J., Chen, V.B., Jain, S., Lewis, S.M., Arendall,

W.B. 3rd, Snoeyink, J., Adams, P.D., Lovell, S.C., Richardson, J.S., Richardson, D.C. MolProbity: More and better reference data for improved all-atom structure validation. *Protein Sci* **27**, 293-315 (2018).

Zhang, G.Y., Campbell, E.A., Minakhin, L., Richter, C., Severinov, K., Darst, S.A. Crystal Structure of Thermus aquaticus Core RNA Polymerase at 3.3 Å Resolution. *Cell* **98**, 811-824 (1999).

Zenkin, N. & Yuzenkova, Y. New Insights into the Functions of Transcription Factors that Bind the RNA Polymerase Secondary Channel. *Biomolecules* **5**, 1195-1209 (2015).

Zhang, K. Gctf: Real-time CTF determination and correction. *J Struct Biol* **193**, 1-12 (2016).

Zheng, S.Q., Palovcak, E., Armache, J.P., Verba, K.A., Cheng, Y., Agard, D.A. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat Methods* **14**, 331-332 (2017).

Zachrdla, M., Padrta, P., Rabatinová, A., Šanderová, H., Barvík, I., Krásný, L., Žídek, L. Solution structure of domain 1.1 of the sigma(A) factor from Bacillus subtilis is preformed for binding to the RNA polymerase core. *J Biol Chem* **292**, 11610-11617 (2017).

7 Appendix

7.1 List of abbreviations

Abbreviation	Expanded
Å	Angstrom (1Å = 10^{-10} m)
AA/aa	amino acid
APS	ammonium peroxodisulfate
ATP	adenosine triphosphate
ADP	adenosine diphosphate
AMP	adenosine monophosphate
AMPPNP	adenylyl-imidodiphosphate
ATPγS	adenosine 5′-O-(3-thio)triphosphate
bp	base pair
BSA	bovine serum albumin
B. s	*Bacillus subtilis*
ßME	2-Mercaptoethanol
°C	degree celsius
cryoEM	electron cryo-microscopy
CTD	C-terminal domain
CTR	C-terminal region
ddH2O	double distilled water
DNA	deoxyribonucleic acid
dsDNA	double-stranded DNA
DTT	1,4-Dithiothreitol
E. coli	*Escherichia coli*
EDTA	ethylenediamine tetraacetic acid, disodium salt dihydrate
EMBL	European molecular biology laboratory
e.g.	exempli gratia
HEPES	4-(hydroxyethyl)-1-piperazineethanesulfonic acid
h	hour
IPTG	Isopropyl-β-D-1-thiogalactopyranoside
kb	kilobase (unit of Na molecule length)
kDa	kilodalton (unit of molecular weight)
l	liter
M	molarity
MALS	multi-angle light scattering
mAU	milli-absorption unit

min	minute
mRNA	messenger RNA
Ni^{2+}-NTA	nickel-nitrilotriacetate
nm	nanometer
NMR	nuclear magnetic resonance
nt	nucleotides
NTD	N-terminal domain
NTP	nucleotide triphosphate
Nus	N-utilization substance
OD	optical density
P	phosphate
PBS	phosphate-buffered saline
PCR	polymerase chain reaction
PDB	protein databank
PEG	polyethylene glycol
rmsd.	root mean square deviation
RNA	ribonucleic acid
RNAP	RNA polymerase
rpm	revolutions per minute
RT	room temperature
s	second
SDS	sodium dodecylsulfate
SDS-PAGE	sodium dodecyl sulfate-polyacrylamide gel electrophoresis
TBE buffer	TRIS-borat-EDTA buffer
TEC	transcription elongation complex
TEMED	N,N,N',N'-tetraethylenediamide
TEV	tobacco etch virus protease
TRIS	TRIS-(hydroxymethyl) aminomethane
T. th	Thermus thermophilus
U	unit
UV	ultraviolet
V	volume
ZnF	Zinc-finger
Nucleic acid bases	
Adenine	A
Cytosine	C
Guanine	G
Thymine	T
Uracil	U

7.2 Complete data of inter-molecular crosslinks in RNAP$^{\Delta\delta\Delta HelD}$-δ-HelD.

Inter-molecular cross-links					
Protein1	Residue1	Protein2	Residue2	Highest Score	Cα-Cα (Å)
HelD	443	β' clamp	314	11.393	7.93
HelD	443	β' clamp	317	7.78	8.45
HelD	37	β' Trigger loop	949	15.601	9.3
HelD	386	β 2-lobe	223	16.669	9.33
HelD	388	β 2-lobe	223	9.172	11.17
HelD	393	β 2-lobe	223	15.204	11.19
HelD	37	β' F-bridge	834	13.462	11.2
HelD	37	β' Trigger loop	948	16.376	12.21
HelD	93	β' jaw	1011	8.759	12.41
HelD	392	β 2-lobe	223	15.132	12.64
HelD	97	β' jaw	1010	11.021	12.68
HelD	536	β 2-lobe	223	13.548	13.06
HelD	37	β' F-bridge	831	8.929	13.26
HelD	537	β 2-lobe	216	11.78	13.56
HelD	442	β' rudder	312	13.34	13.71
HelD	37	β' F-bridge	832	10.791	13.93
HelD	45	β' 2th channel	617	10.918	14.24
HelD	393	β 2-lobe	177	14.001	14.44
HelD	535	β 2-lobe	223	11.126	14.69
HelD	97	β' jaw	1007	10.609	14.75
HelD	433	β' F-bridge α-helix	800	12.841	14.89
HelD	37	β' F-bridge α-helix	837	9.694	15.38
HelD	397	β 2-lobe	190	11.908	15.43
HelD	175	β' jaw	1012	11.727	15.46
HelD	97	β' jaw	1006	12.288	15.51
HelD	396	β 2-lobe	223	11.95	15.76
HelD	397	β 2-lobe	223	14.267	15.76
HelD	97	β' jaw	983	13.815	16.85
HelD	392	β 2-lobe	177	11.621	16.88
HelD	437	β' F-bridge α-helix	793	14.772	17.1
HelD	426	β' jaw	1032	16.241	18.87
HelD	37	β' Trigger loop	939	14.106	19.62
HelD	399	β 2-lobe	223	13.703	20.35
HelD	424	β' jaw	1032	11.964	21

HelD	427	β' jaw	1032	18.178	21.4
HelD	425	β' jaw	1031	13.164	23.49
HelD	425	β' jaw	1033	10.684	24.59
HelD	695	β' jaw	990	13.582	26.59
HelD	428	β' jaw	1033	12.858	27.35
α1	191	β'	520	7.754	25.81
α2	178	β'	649	19.543	12.72
α2	19	β'	520	11.827	14.52
α2	38	β'	649	12.815	14.58
α2	178	β'	648	10.379	15.52
α2	38	β'	650	11.718	16.22
α2	41	β'	649	12.682	16.81
α2	85	β'	545	12.741	20.85
α2	86	β'	545	9.419	22.64
α2	19	β'	541	16.126	22.88
α2	158	β	1014	7.613	45.89
β fork loop	505	β' F-bridge α-helix	785	10.924	9.27
β	1022	β'	649	10.771	12.37
C-terminal β clamp	1111	β' dock	384	13.583	13.27
C-terminal β clamp	1110	β' dock	384	15.442	14.45
β fork loop	518	β' F-bridge α-helix	785	12.544	15.41
C-terminal β clamp	1110	β' dock	382	12.107	17.16
C-terminal β clamp	1109	β' dock	384	15.358	17.94
β	1127	β' clamp	85	7.729	19.26
C-terminal β clamp	1109	β' dock	382	11.437	20.12
C-terminal β clamp	1108	β' dock	407	12.511	22.4
β	1137	β' ZBD	78	11.408	26.34
C-terminal β clamp	1109	β' ZBD	71	12.447	27.19
β flap tip	857	β' dock	384	18.201	29.69
β flap tip	858	β' dock	384	8.263	29.73
β flap tip	858	β' dock	387	10.644	32.04
β flap tip	857	β' dock	387	8.258	32.51
β flap	808	β' dock	384	10.441	34.18
C-terminal β clamp	1111	β' rudder	314	11.345	34.27
β flap tip arms	849	β' dock	382	10.138	34.35
C-terminal β clamp	1108	β'	350	11.635	35.79
β flap	803	β' zipper	40	9.311	37.83
β flap tip arms	867	β' dock	384	14.883	38
β 2-lobe	223	β' jaw	1006	15.319	38.58
β 2-lobe	189	β' Trigger loop	941	12.359	39.61
β flap tip	858	β' clamp	321	11.985	46.73
β flap tip	857	β' clamp	321	16.229	47.06

β connector	924	β'	518	9.903	48.03
β flap	803	β' ZBD	60	13.333	48.67
β	740	β'	742	7.823	48.81
β	739	β'	742	7.946	48.96
β β2-lobe	177	β' jaw	1016	10.508	49.91
β flap tip arms	868	β'	376	11.317	50.27
β flap tip arms	855	β' clamp	110	14.672	59.22
β flap tip	858	β' clamp	108	14.316	59.36
β	932	β' jaw	986	11.756	64.45
β	936	β' jaw	981	8.333	72.44
β	936	β' jaw	986	10.923	73.42
β flap tip arms	868	β' clamp	210	11.561	83.45
β fork loop	557	β' rudder	315	10.944	87.01
β	558	β' rudder	315	10.887	87.17
δ	87	β' jaw	1032	11.341	16.03
δ	85	β' jaw	1032	15.907	16.58
δ	83	β' jaw	1032	12.628	19.34
δ	90	β' jaw	1032	13.574	19.44
Inter-molecular cross-links not defined in the structure					
Protein1	**Residue1**	**Protein2**	**Residue2**	**Highest Score**	**Cα-Cα (Å)**
HelD	516	β'	339	10.046	
HelD	442	β	1065	11.834	
α	294	β flap	804	12.213	
α	294	β	76	10.405	
C-terminal β clamp	1056	β'	334	9.398	
β	1121	β' ZBD	86	9.389	
β flap	881	β'	339	11.084	
β	1127	β' clamp	325	13.555	
β	1137	β' clamp	325	8.717	
β	1115	β'	350	8.996	
β flap tip	857	β'	334	14.005	
β	1121	β' clamp	317	7.885	
β	1126	β' clamp	323	11.687	
β	1121	β'	1	8.785	
β flap tip	557	β' shelf	883	9.549	
C-terminal β clamp	1056	β' F-bridge α-helix	793	15.625	
C-terminal β clamp	1056	β' clamp	322	15.063	
C-terminal β clamp	1067	β'	350	11.394	
C-terminal β clamp	1056	β' shelf	900	11.773	
β	1127	β' clamp	324	11.754	
β flap	881	β'	338	12.717	
β	1161	β' ZBD	71	13.008	

C-terminal β clamp	1066	β'	611	13.145	
C-terminal β clamp	1067	β' F-bridge α-helix	799	14.435	
C-terminal β clamp	1055	β' jaw	1032	14.444	
C-terminal β clamp	1053	β' jaw	1032	14.733	
C-terminal β clamp	1056	β'	384	10.492	
C-terminal β clamp	1056	β' F-bridge α-helix	792	10.531	
β	563	β'	393	10.682	
C-terminal β clamp	1056	β' clamp	323	10.766	
C-terminal β clamp	1069	β'	339	10.775	
β	1115	β'	422	17.623	
β	1115	β'	426	18.904	
α1/2	3	α2/1	91	8.009	
α1/2	1	α2/1	93	8.38	
α1/2	1	α2/1	92	8.706	
Intra-molecular cross-links					
Protein1	**Residue1**	**Protein2**	**Residue2**	**Highest Score**	**Cα-Cα (Å)**
HelD	395	HelD	398	12.725	5
HelD	371	HelD	373	11.54	5.41
HelD	321	HelD	323	13.732	5.51
HelD	449	HelD	453	8.477	5.9
HelD	238	HelD	240	9.907	5.91
HelD	212	HelD	216	6.549	6.16
HelD	640	HelD	642	12.66	6.39
HelD	424	HelD	428	9.455	7.12
HelD	647	HelD	652	10.785	8.37
HelD	639	HelD	642	11.781	9.17
HelD	32	HelD	38	11.741	10.04
HelD	212	HelD	219	8.805	10.24
HelD	655	HelD	662	13.612	10.42
HelD	32	HelD	39	7.176	10.55
HelD	225	HelD	599	8.503	11.41
HelD	654	HelD	662	8.132	12.37
HelD	368	HelD	375	10.236	12.59
HelD	150	HelD	155	22.729	13.22
HelD	498	HelD	507	7.773	14.52
HelD	652	HelD	662	11.302	15.06
HelD	690	HelD	695	7.961	15.81
HelD	695	HelD	703	10.133	16.71
HelD	435	HelD	462	9.085	17.21
HelD	438	HelD	462	7.061	20.27
HelD	354	HelD	361	8.925	20.63
α	155	α	158	13.72	8.81

α1	204	α1	208	14.285	6.4
α1	191	α1	194	12.558	9.53
α1	18	α1	22	12.669	10.69
α1	19	α1	84	10.287	15.92
α2	15	α2	20	13.4	8.54
β flap	802	β flap	804	13.404	5.2
β gate loop	229	β gate loop	232	15.476	5.22
β	81	β	84	15.241	5.24
β	81	β	86	15.222	5.26
β gate loop	251	β gate loop	253	14.436	5.4
β 1-lobe	82	β 1-lobe	85	16.709	5.4
β 1-lobe	81	β 1-lobe	83	10.701	5.7
β	139	β	141	8.328	6.12
β gate loop	229	β gate loop	233	13.502	6.15
β	924	C-terminal β clamp	1043	6.076	6.31
β 2-lobe	263	β 2-lobe	266	8.888	6.33
β	708	β	710	8.731	6.64
β fork loop	556	β	584	11.597	7.15
β 2-lobe	174	β 2-lobe	178	10.879	7.21
β connector	924	C-terminal β clamp	1041	6.344	7.37
β	697	β	704	8.991	7.62
β SI1	299	β SI1	307	11.765	7.9
β	2	β	1002	13.658	8.6
β fork loop	554	β	585	10.582	8.65
β flap	876	β flap	879	8.815	9.12
β flap tip arms	855	β flap tip	860	10.812	9.15
β	702	β	705	12.955	9.3
β	147	β	150	10.295	9.48
β gate loop	229	β gate loop	235	12.533	9.91
β connector	909	β connector	912	8.453	9.99
β 2-lobe	222	β 2-lobe	226	11.58	10.18
β flap tip arms	855	β flap tip	861	8.77	10.26
β flap tip arms	849	β flap tip arms	867	11.188	10.41
β 1-lobe	83	β 1-lobe	87	8.442	10.46
β fork loop	554	β	584	11.436	10.72
β flap tip arms	849	β flap tip	862	15.884	10.79
β 1-lobe	83	β 1-lobe	88	9.72	10.84
β flap	803	β flap tip arms	851	10.868	10.84
β	698	β	704	10.214	10.93
β 2-lobe	156	β 2-lobe	178	10.884	11.13
β 2-lobe	177	β 2-lobe	222	14.343	11.13
β 2-lobe	190	β 2-lobe	223	21.295	11.23

β flap	876	β flap	880	10.23	11.48
β	702	β	706	16.372	11.68
β	731	β	741	9.886	12.03
β 2-lobe	189	β 2-lobe	224	12.807	12.23
β	683	β	704	8.951	12.42
β 2-lobe	177	β 2-lobe	224	14.935	12.53
β 2-lobe	189	β 2-lobe	223	9.454	12.97
β	731	β	740	9.008	13
β	799	β	803	13.773	13.02
β flap tip arms	849	β flap tip arms	866	11.049	13.2
β connector	797	β flap	801	11.603	13.43
β flap	891	β flap	900	9.308	13.97
β SI1	282	β	571	9.288	14.22
β flap	891	β flap	897	9.476	14.28
β flap	876	β flap	881	8.786	14.53
β 1-lobe	76	β	709	11.636	15
β 1-lobe	80	β	709	14.161	15.1
β 1-lobe	223	β gate loop	235	13.254	15.36
β SI1	282	β	568	12.613	15.65
β SI1	283	β	572	7.816	15.69
β 2-lobe	189	β 2-lobe	222	17.16	16.36
β	149	β 2-lobe	189	13.363	16.49
β	702	β	708	9.91	16.79
β flap	803	β flap tip	862	11.178	17.08
β SI1	282	β	572	19.705	17.51
β 2-lobe	189	β 2-lobe	227	14.316	17.6
β	702	β	735	12.264	17.69
β	521	β	932	13.007	17.77
β	702	β	736	8.396	18.12
β 2-lobe	223	β gate loop	234	12.069	18.45
β SI1	282	β	566	8.934	18.61
β	802	β flap tip	862	12.416	18.62
β flap	876	β flap	885	9.144	18.75
β 1-lobe	83	β	709	9.02	18.8
β 1-lobe	80	β	677	8.603	20.23
β SI1	281	β	572	10.659	20.78
β 2-lobe	156	β 2-lobe	212	13.178	22.35
β	1001	β	1018	8.271	24.93
C-terminal β clamp	1108	β	1135	9.999	26.01
β	610	β	739	8.485	31.94
β flap tip	857	B	1108	8.941	33.45
β	610	β	740	7.207	34.23

β	610	β	741	7.214	35.81
β protrusion	461	β	634	7.462	37.2
β	601	β	740	8.113	44.3
β	604	β	740	7.869	45.09
β	558	β flap	799	8.652	70.01
β fork loop	556	β flap	799	8.652	75.18
β 2-lobe	224	β flap	876	8.181	78.58
β' jaw	1010	β' jaw	1025	14.2	4.17
β' jaw	1011	β' jaw	1025	10.414	4.46
β' ZBD	71	β' ZBD	83	11.512	4.61
β' jaw	1011	β' jaw	1024	12.418	5.08
β' clamp	195	β' clamp	197	14.697	5.37
β' F-bridge α-helix	792	β' F-bridge α-helix	794	13.164	5.41
β' clamp	160	β' clamp	162	14.775	5.48
β'	582	β'	584	13.498	5.52
β' jaw	1011	β' jaw	1026	16.207	5.66
β' jaw	1124	β' jaw	1126	14.27	5.81
β' clamp	106	β' clamp	109	16.568	5.83
β' zipper	28	β' zipper	30	12.181	5.87
β' jaw	984	β' jaw	987	11.952	6.17
β'	706	β'	710	9.186	6.28
β' rudder	313	β' rudder	315	9.337	6.46
β'	588	β'	590	12.453	6.51
β' ZBD	70	β' ZBD	83	17.427	6.51
β' F-bridge α-helix	785	β' Trigger loop	940	10.151	6.63
β' ZBD	55	β' ZBD	57	11.299	7.02
β'	582	β'	586	9.077	7.03
β' dock	403	β' dock	407	14.325	7.13
β'	543	β'	546	12.421	7.15
β'	353	β'	427	11.771	7.46
β'	519	β'	544	14.765	7.5
β'	516	β'	519	20.294	7.58
β' ZBD	71	β' ZBD	81	13.146	7.77
β'	586	β'	622	11.437	7.86
β'	350	β'	353	11.863	7.94
N-terminal β‘clamp	162	N-terminal β‘clamp	166	12.473	8.34
β' F-bridge α-helix	793	β' Trigger loop	939	13.604	8.38
β' zipper	30	β' zipper	43	10.132	8.48
β' clamp	153	β' clamp	158	9.506	8.49
β' F-bridge α-helix	790	β' Trigger loop	940	9.082	8.53
β' ZBD	55	β' ZBD	61	7.24	8.59
β' ZBD	55	β' ZBD	58	8.121	8.61

β' clamp	191	β' clamp	196	11.33	8.61
β' jaw	1031	β' jaw	1036	13.857	8.61
β' F-bridge	820	β' F-bridge	823	8.634	8.64
β' F-bridge α-helix	790	β' F-bridge α-helix	795	13.681	8.64
β' rudder	312	β' rudder	315	18.017	8.88
β'	349	β'	644	9.027	9.3
β' jaw	977	β' jaw	991	10.39	9.44
β'	352	β'	638	10.332	9.46
β' dock	403	β' dock	408	9.597	9.59
β'	587	β'	622	15.015	9.78
β'	552	β'	556	12.444	9.81
β'	617	β'	623	12.262	9.85
β' zipper	28	β' zipper	43	9.173	9.98
β' F-bridge	829	β' F-bridge	833	13.45	10.47
β'	348	β'	649	12.464	10.47
β' zipper	43	β' zipper	51	5.961	10.5
β'	616	β'	622	13.138	10.52
β'	672	β'	715	9.039	10.81
β'	781	β' F-bridge α-helix	785	9.181	10.87
β' zipper	43	β' ZBD	61	13.165	10.88
β' ZBD	71	β' ZBD	85	8.635	10.92
β'	518	β'	545	13.698	10.92
β'	615	β'	622	10.331	11.02
β'	352	β'	636	8.744	11.2
β'	349	β'	649	13.751	11.22
β'	616	β'	623	10.563	11.22
β' clamp	185	β' clamp	192	13.583	11.29
N-terminal β‘clamp	161	N-terminal β‘clamp	166	9.948	11.35
N-terminal β‘clamp	173	β' clamp helices	285	10.749	11.44
N-terminal β‘clamp	169	β' clamp helices	282	10.652	11.5
β'	342	β'	358	8.33	11.68
β' zipper	40	β' zipper	44	10.673	11.82
β' F-bridge	820	β' F-bridge	824	8.78	12.1
β' F-bridge	827	β' F-bridge	831	6.794	12.19
β' clamp	1154	β' clamp	1158	14.735	12.25
β'	367	β'	375	8.118	12.32
β'	616	β'	727	11.646	12.34
β'	621	β'	629	8.673	12.36
β'	353	β'	638	13.091	12.41
β' clamp	1152	β' clamp	1156	13.311	12.52
β'	514	β'	519	9.942	12.53
β' zipper	40	β' ZBD	59	12.181	13.4

β'	831	β' jaw	1011	11.689	13.96
N-terminal β‘clamp '	169	β' clamp helices	285	9.364	14.04
β'	615	β'	727	10.727	14.11
β'	350	β'	649	10.895	14.51
β' dock	358	β' dock	367	12.355	14.88
β'	569	β'	649	10.985	15.01
N-terminal β‘clamp	175	β' clamp helices	285	8.592	15.05
β'	526	β'	649	8.816	15.05
β' dock	377	β' dock	407	14.36	15.17
β' ZBD	70	β' ZBD	78	13.679	15.24
β' zipper	43	β' ZBD	59	8.679	15.35
β' ZBD	71	β' ZBD	78	11.084	15.77
β'	668	β'	678	9.393	15.8
β'	482	β'	492	10.668	16.2
β' jaw	1030	β' jaw	1125	9.527	17.08
β' dock	377	β' dock	408	8.041	17.14
β' zipper	44	β' ZBD	56	8.846	17.36
β' jaw	1032	β' jaw	1127	14.803	17.59
β' clamp	1152	β' clamp	1158	10.299	17.61
β' zipper	29	β' zipper	40	6.727	17.8
N-terminal β‘clamp	168	β'	285	12.045	17.82
β' zipper	43	β' ZBD	57	9.143	18.64
β' jaw	1032	β' jaw	1129	13.24	18.86
β' jaw	1028	β' jaw	1123	10.806	19.28
β'	661	β'	670	13.069	19.47
β' jaw	1029	β' jaw	1124	11.595	20.13
β'	830	β' jaw	1010	11.326	20.37
β' clamp	1152	β' clamp	1159	9.519	20.58
β' zipper	36	β' ZBD	78	8.36	21.18
β'	661	β'	671	10.885	21.92
β'	357	β'	637	8.424	22.08
β' jaw	1006	β' jaw	1034	13.896	22.86
β'	661	β'	672	10.836	22.95
β' jaw	1006	β' jaw	1032	14.846	24.19
β'	357	β'	636	7.876	25.77
β'	9	β' ZBD	78	14.866	38.68
β' clamp	187	β' F-bridge α-helix	790	5.693	59.74
β'	518	β' F-bridge α-helix	799	6.245	62.36
β' ZBD	76	β'	519	7.299	90.27
Intra-molecular cross-links not defined in the structure					
Protein1	**Residue1**	**Protein2**	**Residue2**	**Highest Score**	**Cα-Cα (Å)**
HelD	455	HelD	462	10.813	

HelD	358	HelD	360	8.069	
α	279	α	281	11.967	
α	278	α	281	8.915	
C-terminal β clamp	1060	C-terminal β clamp	1068	13.565	
β	1182	β	1188	11.58	
β SI1	299	β SI1	311	7.268	
β SI1	299	β SI1	309	7.335	
β	1146	β	1167	6.477	
C-terminal β clamp	1046	C-terminal β clamp	1056	5.867	
C-terminal β clamp	1044	C-terminal β clamp	1056	6.747	
β	1165	β	1182	10.232	
β	1046	C-terminal β clamp	1053	13.872	
β	1146	β	1159	9.919	
β	1146	β	1165	9.553	
C-terminal β clamp	1044	C-terminal β clamp	1053	7.064	
C-terminal β clamp	1110	β	1116	9.82	
C-terminal β clamp	1044	C-terminal β clamp	1067	8.26	
β flap tip	860	β flap tip arms	868	8.412	
β flap tip arms	855	C-terminal β clamp	1069	8.597	
β flap tip arms	853	C-terminal β clamp	1067	8.86	
β flap	881	C-terminal β clamp	1067	9.087	
β 2-lobe	219	β	1182	14.077	
β 2-lobe	220	β	1186	14.358	
β	1146	β	1158	9.469	
β	1146	β	1156	10.664	
β flap	879	C-terminal β clamp	1053	9.489	
β flap	876	C-terminal β clamp	1070	9.559	
β flap	880	C-terminal β clamp	1067	9.944	
β connector	795	C-terminal β clamp	1049	10.981	
β flap tip arms	849	β flap tip arms	849	11.621	
β flap tip	857	C-terminal β clamp	1067	13.976	
β	1146	β	1156	10.651	
β flap tip	857	C-terminal β clamp	1052	18.524	
β' C-term	1161	β' C-term	1169	6.744	
β' C-term	1168	β' C-term	1175	8.795	
β' jaw	1024	β' jaw	1032	11.701	
β' C-term	1162	β' C-term	1170	10.721	
β' clamp	193	β' clamp	197	8.268	
β' clamp	330	β' clamp	1154	11.815	
β'	1	β' C-term	1166	8.182	
β'	419	β' F-bridge α-helix	793	8.37	
β' clamp	1154	β' C-term	1164	8.664	

β' clamp	1154	β' C-term	1163	15.731	
β' jaw	1010	β' jaw	1017	10.951	
β' clamp	329	β' clamp	1154	11.201	
β'	340	β'	342	10.018	
β'	544	β'	545	12.661	
β'	417	β'	421	14.04	
β'	417	β'	423	10.305	
β' C-term	1159	β' C-term	1162	10.705	

8 Curriculum vitae

For reasons of data protection, the curriculum vitae is not published in the electronic version.