The Core and Holoenzyme Forms of RNA Polymerase from *Mycobacterium smegmatis*

Tomáš Kouba,^{a*} Jiří Pospíšil,^{b,c} Jarmila Hnilicová,^b Hana Šanderová,^b Ivan Barvík,^d Libor Krásný^b

^aMRC Laboratory of Molecular Biology, Cambridge, United Kingdom

^bLaboratory of Microbial Genetics and Gene Expression, Institute of Microbiology, Academy of Sciences of the Czech Republic, Prague, Czech Republic

^cDepartment of Genetics and Microbiology, Faculty of Science, Charles University in Prague, Prague, Czech Republic

^dDivision of Biomolecular Physics, Institute of Physics, Faculty of Mathematics and Physics, Charles University in Prague, Prague, Czech Republic

ABSTRACT Bacterial RNA polymerase (RNAP) is essential for gene expression and as such is a valid drug target. Hence, it is imperative to know its structure and dynamics. Here, we present two as-yet-unreported forms of *Mycobacterium smegmatis* RNAP: core and holoenzyme containing σ^A but no other factors. Each form was detected by cryo-electron microscopy in two major conformations. Comparisons of these structures with known structures of other RNAPs reveal a high degree of conformational flexibility of the mycobacterial enzyme and confirm that region 1.1 of σ^A is directed into the primary channel of RNAP. Taken together, we describe the conformational changes of unrestrained mycobacterial RNAP.

IMPORTANCE We describe here three-dimensional structures of core and holoenzyme forms of mycobacterial RNA polymerase (RNAP) solved by cryo-electron microscopy. These structures fill the thus-far-empty spots in the gallery of the pivotal forms of mycobacterial RNAP and illuminate the extent of conformational dynamics of this enzyme. The presented findings may facilitate future designs of antimycobacterial drugs targeting RNAP.

KEYWORDS RNA polymerase, bacterial transcription, conformational change, cryo-electron microscopy, mycobacteria, protein structure, transcription initiation factor

Mycobacteria are a group of medically important bacteria with a number of species that cause serious human and animal diseases, such as tuberculosis, which ranks among the top 10 causes of death worldwide according to the World Health Organization. Understanding the molecular machinery of these organisms is necessary to combat these pathogens.

RNA polymerase (RNAP) is a central enzyme of gene expression and, as such, it is an attractive drug target. Rifampin, an antibiotic binding to RNAP, has been used to treat tuberculosis since the late 1960s, but mycobacteria are able to gain resistance due to point mutations of amino acids in the rifampin-binding cavity (1). Detailed knowledge of the RNAP structure and dynamics is thus fundamental for the rational design of new RNAP inhibitors (2).

RNAP is a multisubunit enzyme consisting of an α dimer and β , β' , and ω subunits (3). This RNAP core is catalytically active, but to initiate transcription it needs a sigma factor. σ^A is the principal sigma factor responsible for transcription of housekeeping genes (4). Eventually, other subunits, in particular, various transcription activators such as CarD (5) or RpbA (6, 7), may be associated with the RNAP holoenzyme.

σ^A (σ^{70} in *Escherichia coli*) is composed of domains 1 to 4, some of which are further subdivided (4). Interestingly, region 1.1 of σ^A is still somewhat elusive even after almost two decades of structural studies. It is usually partly or fully missing in the bacterial RNAP crystal structures. A notable exception is $\sigma^{A_{1.1}}$ from *E. coli*, which was solved by

Citation Kouba T, Pospíšil J, Hnilicová J, Šanderová H, Barvík I, Krásný L. 2019. The core and holoenzyme forms of RNA polymerase from *Mycobacterium smegmatis*. *J Bacteriol* 201:e00583-18. <https://doi.org/10.1128/JB.00583-18>.

Editor Ann M. Stock, Rutgers University-Robert Wood Johnson Medical School

Copyright © 2019 American Society for Microbiology. All Rights Reserved.

Address correspondence to Tomáš Kouba, tkouba@embl.fr, or Libor Krásný, krasny@biomed.cas.cz.

* Present address: Tomáš Kouba, EMBL Grenoble, Grenoble, France.

Received 20 September 2018

Accepted 20 November 2018

Accepted manuscript posted online 26 November 2018

Published 28 January 2019

crystallography in complex with RNAP, where it is positioned in the primary channel of RNAP (8), from which it must be displaced by DNA during transcription initiation. Otherwise, only isolated structures of $\sigma^{A}_{1.1}$ from *Thermotoga maritima* and *Bacillus subtilis* were determined by using nuclear magnetic resonance (9, 10).

The 1.1 domains from all these organisms are playing the same roles. They prevent binding of the σ^A free form to promoter DNA. Further, they prevent access of DNA to the active site of RNAP prior to formation of the catalytically competent transcription initiation complex. However, the structural details between 1.1 domains from different organisms vary widely. Currently, three types of known structures of this domain exist: (i) the highly similar structures, as solved for *B. subtilis* and *E. coli*, consisting of three helices (HI to HIII) packed in antiparallel manner (8, 10); (ii) the structure from *T. maritima*, where HI packs perpendicularly to HII and HIII (9); and (iii) the possibly disordered and flexible structure from mycobacteria with only the $\sigma^A_{N\text{-helix}}$ discernible (11–14).

Transcription initiation is a highly dynamic process and conformational changes of RNAP are required for it to happen (15, 16). RNAP first forms the closed complex (RP_C) with promoter DNA, in which the two DNA strands are still together; subsequently, the complex isomerizes through a kinetic intermediate(s) to the open complex (RP_O), where the DNA is opened from positions -10 to $+2$ (relative to $+1$, the transcription start) (17–19). In mycobacteria, the conformational change from RP_C to RP_O is greatly stimulated by the presence of transcriptional factors RpbA and CarD in a cooperative manner (12, 20). RP_O is also stabilized by the presence of nucleoside triphosphates (NTPs) complementary to the $+1$ and, to a lesser degree, the $+2$ transcription positions of the template strand (10, 21).

Historically, the established structural models of bacterial transcription were based on studies of RNAP from *Thermus aquaticus*, *Thermus thermophilus*, and *Escherichia coli*. However, even alignments of amino acids of individual subunits from various species reveal significant differences between these enzymes and studies addressing this variety have now only just begun (22). Hence, structural studies were recently published, focusing on specific transcription systems, especially those from, or related to pathogenic species, such as *Mycobacterium tuberculosis* (11, 14, 23) and *Mycobacterium smegmatis* (12, 13).

Currently, cryo-electron microscopy (cryo-EM) structures are available for *M. tuberculosis* RNAP σ^A in complexes (see Table 1) with RbpA (PDB entry 6C05 [11]), fidaxomicin (PDB entry 6FBV [23]), or both (PDB entry 6C06 [11]) and with RbpA and double-fork DNA (PDB entry 6C04 [11]) or RbpA, fidaxomicin, and upstream-fork DNA (PDB entry 6BZO [11]). Moreover, crystal structures have been reported for *M. smegmatis* RNAP σ^A in complex with RbpA and upstream fork DNA (PDB entry 5V18 [12]), RP_O with RbpA (PDB entry 5V15 [13]), the *M. tuberculosis* RNAP σ^A holoenzyme with downstream fork DNA and with various inhibitors (including rifampin and *N*- α -aroyl-*N*-aryl-phenyl-alaninamides) and short fragments (2 to 4 nucleotides [nt]) of RNA (PDB entries 5UHA, 5UHB, 5UHE, 5UHF, 5UHG, 5UH7, 5UH9, 5UH5, 5UH8, 5UH6, 5UHC, and 5UHD [14]) (see Table 1).

Compared to previously known structures of RNAPs from other organisms, the mycobacterial RNAPs contain a prominent feature, a long α -helical coiled-coil domain protruding from the β' clamp (β' il), and the models also depict the position of a part of the N-terminal segment of the σ^A factor. Both of these elements are involved in specific mechanisms of transcription initiation. It was proposed that the taxon-specific mycobacterial RNAP structural features might help stabilize the very labile mycobacterial RP_O (14), but the mechanistic details are still debated (12). Furthermore, the cryo-EM three-dimensional (3D) structures, which are captured in almost native states, hint at the importance of the β' clamp dynamics during transition from RP_C to RP_O .

Here, using cryo-EM, we describe the as-yet-unreported *M. smegmatis* RNAP core and holoenzyme structures containing σ^A but no other factors. The structural information presented here provides further details on the taxon-specific nonconserved extension, $\beta'i1$, in the β' subunit of RNAP, as well as the N-terminal part of the σ^A factor

(region 1.1). In summary, the RNAP core and RNAP σ^A structures, combined with the published models of *M. smegmatis* and *M. tuberculosis* RNAPs (see Tables 1 and 2), reveal the extent of conformational changes of mycobacterial RNAP and complete the gallery of the pivotal forms of the enzyme.

RESULTS AND DISCUSSION

Cryo-EM microscopy of *M. smegmatis* RNAP. We cloned all the *M. smegmatis* RNAP core subunits into a plasmid (see Fig. S1A in the supplemental material), overexpressed in *E. coli*, and purified. The purification procedure consisted of affinity tag purification, followed by high-salt washes and size exclusion chromatography (Fig. S1B). The elution peak included a fractional shoulder, suggesting a minor heterogeneity of the sample. We also found that intermediate ion exchange cleaning steps induced a loss of the ω subunit and therefore were excluded from the purification procedure. Also, we prepared *M. smegmatis* σ^A (for details, see Materials and Methods). Subsequently, we used these recombinant proteins to reconstitute the *M. smegmatis* RNAP σ^A holoenzyme (Fig. S1B). We verified its enzymatic activity in an *in vitro* transcription assay using a DNA template with a native ribosomal promoter (Fig. S1C).

Several vitrification conditions were tested to prepare cryo-EM specimens of the *M. smegmatis* RNAP σ^A holoenzyme. Free standing vitreous ice conditions introduced a strong preferred orientation of RNAP molecules and were limited to a single view, and, consequently, unsuited for 3D reconstructions. To circumvent this problem, we used graphene oxide supports, which allowed a much broader distribution of RNAP orientations (for details see Materials and Methods and Fig. S2A, B, and E). However, the graphene oxide support required a relatively low protein concentration (~ 110 nM) in the vitreous specimen, and the low protein concentration caused a partial dissociation of the RNAP holoenzyme (σ^A binds to the RNAP core with a K_d of ~ 90 nM, as measured by isothermal titration calorimetry [data not shown]).

In total, we imaged $\sim 500,000$ single-particle images of RNAP molecules. Several rounds of unsupervised 3D particle classifications to high degree of angular accuracy (for details, see Materials and Methods) separated all presented structural species and produced cryo-EM densities for each of them (Fig. S2C and D). Primarily, we targeted the *M. smegmatis* RNAP σ^A holoenzyme, but as a result of the low protein concentration in the cryo-EM specimens, we also identified RNAP core particles in the data set.

Hence, two main forms of *M. smegmatis* RNAP were identified: core and holoenzyme. Moreover, both of them were detected in two major conformations. These conformations were represented by the two most abundant 3D classes of each form, and they were clearly distinguishable by the correlation coefficient of fit in the map tool in Chimera (for the number of particles in each class, see Fig. S2C). The remaining minor 3D classes were not analyzed. The resolutions of the cryo-EM density maps for the most abundant 3D classes of the *M. smegmatis* RNAP core (here referred to as Core1) and RNAP σ^A holoenzyme (here referred to as Holo1) were estimated at 3.8 and 4.2 Å at the 0.143 gold standard Fourier shell correlation (FSC) cutoff (Fig. S2F). The resolutions of the cryo-EM density maps for the second most abundant 3D classes, here referred to as Core2 and Holo2, were estimated at 4.3 and 4.9 Å, respectively (Fig. S2F). Local resolution calculations (Fig. S3) indicated that the central part of the complexes was more defined than the peripheries. A particle orientation efficiency score (24) of ~ 0.47 (Table S1) suggested a minor Fourier space gap in the 3D electron densities volumes and probably slightly hampered the final resolution (a particle orientation efficiency score of >0.6 is reported to be optimal [18]), but no 3D deformation related to poor particle distribution was observed.

Construction of atomic models of *M. smegmatis* RNAP. The protein chains of *M. smegmatis* RNAP were fitted into the produced cryo-EM densities. Atomic models of core and holoenzyme were generated (Fig. 1; see also Fig. S3 and S4) using the crystal structure of *M. smegmatis* transcription initiation complex (TIC) and RP_{O} as the templates (PDB entries 5VI8 and 5VI5 [12, 13]). For further details, see Materials and Methods.

FIG 1 Cryo-EM structure determinations of *M. smegmatis* RNAP complexes. (Top row) Secondary structures of two forms of RNAP from *M. smegmatis* as observed in the experimental cryo-EM densities: core (left) and holoenzyme (right, RNAP in complex with primary sigma factor, σ^A). Subunits are colored as follows: β , yellow; β' , green; α , gray; α' , cyan; ω , orange; σ^A , magenta. (Middle row) Cryo-EM maps are filtered and colored according to local resolution calculated with the RELION software package (see Fig. S3 for more detail). (Bottom row) Cryo-EM maps, together with fit of secondary structure models of *M. smegmatis* RNAP complexes. Maps are visualized as 0.5 transparent at 0.08 volume threshold in Chimera (see Fig. S4 for more detail). Color coding is as described for the top row.

Next, we superimposed the obtained models over the other recent X-ray and cryo-EM structures and this provided three main insights into the structure of *M. smegmatis* RNAP: (i) structural arrangement of the primary channel, the binding site for the downstream DNA (dwDNA) in the RNAP core; (ii) movements of the nonconserved coiled-coil domain β' i1 that is specific for mycobacteria; and (iii) interaction of the RNAP core with σ^A , including the position of the structured part of the σ^A region 1.1 in the holoenzyme. This information contributes to the recent discussion of how these two latter elements influence the transcription initiation and especially the formation of RP_{O_0} , as they virtually block passage of DNA template into the active site of RNAP.

(i) Conformational changes of conserved parts of the *M. smegmatis* RNAP primary channel. The overall shape of RNAP resembles a crab claw with the two pincers (the clamp consisting of parts of β and β' subunits) forming the primary channel. In the primary channel, RNAP interacts with 14 bp of the DNA downstream of the active site (25). This interaction causes changes in the positions of the two pincers along the defined molecular hinges of switch regions (Table S2) (26). It was shown that the RNAP clamp is predominantly open in the free RNAP core form and RP_{C_0} , but closed

TABLE 1 Distances between the β and β' subunits at the narrowest point of the dwDNA entrance: β Lys271C α and β' Lys123C α (*M. smegmatis*) and β Lys280C α and β' Lys123C α (*M. tuberculosis*)^a

RNAP complex	Organism	Distance (Å)	Residues	PDB entry	Source or reference
Core1	<i>M. smegmatis</i>	28.0	K271-K123	6F6W ^{EM}	This study
Core2	<i>M. smegmatis</i>	32.2	K271-K123	NA	This study
Holo1	<i>M. smegmatis</i>	29.4	K271-K123	6EYD ^{EM}	This study
Holo2	<i>M. smegmatis</i>	33.6	K271-K123	NA	This study
holo + RbpA	<i>M. tuberculosis</i>	29.9	K280-K123	6C05 ^{EM}	11
holo + LpM	<i>M. tuberculosis</i>	34.2	K280-K123	6FBV ^{EM}	23
holo + RbpA + Fdx	<i>M. tuberculosis</i>	33.6	K280-K123	6C06 ^{EM}	11
holo + RbpA + Fdx + up	<i>M. tuberculosis</i>	32.3	K280-K123	6BZO ^{EM}	11
holo + RbpA + up + down	<i>M. tuberculosis</i>	19	K280-K123	6C04 ^{EM}	11
holo + down	<i>M. tuberculosis</i>	18.7	K280-K123	5UHA ^{XR}	14
holo + down + Rif	<i>M. tuberculosis</i>	18.8	K280-K123	5UHB ^{XR}	14
holo + down + D-AAP1	<i>M. tuberculosis</i>	18.7	K280-K123	5UHE ^{XR}	14
holo + down + D-IX336	<i>M. tuberculosis</i>	18.3	K280-K123	5UHF ^{XR}	14
holo + down + Rif + D-AAP1	<i>M. tuberculosis</i>	18.3	K280-K123	5UHG ^{XR}	14
holo + RbpA + up	<i>M. smegmatis</i>	17.8	K271-K123	5VI8 ^{XR}	12
RP _o + RbpA	<i>M. smegmatis</i>	20.0	K271-K123	5VI5 ^{XR}	13
holo + down + 2 nt RNA	<i>M. tuberculosis</i>	18.3	K280-K123	5UH9 ^{XR}	14
holo + down + 3 nt RNA	<i>M. tuberculosis</i>	18.5	K280-K123	5UHS ^{XR}	14
holo + down + 4 nt RNA	<i>M. tuberculosis</i>	18.4	K280-K123	5UH8 ^{XR}	14
holo + down + Rif + 2 nt RNA	<i>M. tuberculosis</i>	18.7	K280-K123	5UH6 ^{XR}	14
holo + down + Rif + 3 nt RNA	<i>M. tuberculosis</i>	18.2	K280-K123	5UHC ^{XR}	14
holo + down + Rif + 4 nt RNA	<i>M. tuberculosis</i>	18.5	K280-K123	5UHD ^{XR}	14

^a“up” stands for DNA upstream fork; “down” stands for DNA downstream fork. nt, nucleotide(s); holo, holoenzyme; Fdx, fidaxomicin; LpM, lipiarmycin A3; Rif, rifampin; D-AAP1, *N*- α -(benzenecarbonyl)-*N*-(2-methylphenyl)-*D*-phenylalaninamide; D-IX336, *N*-(2-methylphenyl)-*N*- α -(selenophene-2-carbonyl)-*D*-phenylalaninamide; NA, not applicable. Superscripts: XR, X-ray; EM, Cryo-EM.

in RP_o, initial transcribing complexes (RP_{TC}), and a transcription elongation complex (EC) (15, 27).

We quantitated the extent of these conformational changes (Table 1 and Fig. 2; see also Fig. S5) between various forms of the mycobacterial enzyme by measuring the distance between the β and β' subunits at the narrowest point of the dwDNA entrance (i.e., distance between β Lys271C α and β' Lys123C α). This distance is only approximately 17.8 to 20 Å (Table 1) in the closed dwDNA primary channel-occupied structures, as well as in structures containing RbpA and upstream fork DNA 5VI8 (12) that securely hold DNA. On the other hand, it is approximately 28.0 to 34.2 Å (Table 1) in the case of the open core and holoenzyme structures, which would allow a double-stranded dwDNA to enter the primary channel. The open conformation seems to be an intrinsic feature of the RNAP core itself, as observed for both the Core1 and Core2, and is not dependent on the presence the σ factor, a finding consistent with previous observations (15). We also confirm that the mycobacterial holoenzyme favors the open state (11, 15, 23, 27), since the Holo1 conformation is similar to the relaxed open state of the RNAP Holo/RbpA (11), and the Holo2 conformation is similar to the one in the extreme open form trapped by fidaxomicin (11, 23). This in agreement with the observation that the conformational change toward the close state occurs only after binding of dwDNA.

(ii) RNAP core and holoenzyme structures reveal a new position of the non-conserved domain β' i1 of *M. smegmatis* RNAP. Even though the local resolution of β' i1 is relatively low (Fig. S4), we were able to localize its position according to the fit of secondary structures (Fig. S5 and S6). β' i1 was previously reported to hold its extended coiled-coil fold independently (14), and thus even the incomplete density may determine its full position. However, the lack of cryo-EM density in the area where the tip of the domain would be predicted suggests that this region displays some degree of flexibility relative to the rest of the domain.

Contrary to the previously characterized X-ray and cryo-EM closed-clamp structures containing dwDNA (Table 1) (11–14), but similar to the recent open-clamp holoenzyme complexes with RbpA and/or fidaxomicin determined by cryo-EM (Table 1) (11, 23), the position of the β' i1 domain in both our open-clamp RNAP core and holoenzyme structures is markedly swung outwards relative to the β domain 2 (Fig. 2). This

FIG 2 *M. smegmatis* RNAP exhibits conformational dynamics. Conformational dynamics are illustrated by distances (arrows) between the β and β' subunits at the narrowest point of the dwDNA entrance (distance between β Gly271 α and β' Lys123 α) and by positioning of the nonconserved domain β' i1 (distances between β Leu266 α and β' Arg214 α) in the main conformation of *M. smegmatis* RNAP. Core1/Core2 and Holo1/Holo2 represent the two most represented conformations of respective forms of *M. smegmatis* RNAP. Closed RNAP conformations of *M. smegmatis* RPo (PDB entry 5VI5) and *M. tuberculosis* holoenzyme in complex with the antibacterial compound fidaxomicin (PDB entry 6FBV) that traps the open conformation of the pincer are shown for comparison. RNAP subunits and nucleic acids chains are colored as follows: β , yellow; β' , green; α , gray; α' , cyan; ω , orange; σ A, magenta; DNA duplex, blue and cyan.

displacement increases further from the rotational axis defined by the switch regions (Table S2), and the central part of the β' i1, which is the closest to the β domain 2 (Table S2), is relocated up to ~ 25 Å (in terms of the β Leu266 α – β' Arg214 α distance) (Table 2 and Fig. 2; see also Fig. S5 and S6). This conformation also widens the passage between β' i1 and β domain 2 to around 30 Å and at the same time promotes a steric funneling toward the primary channel.

Lin et al. (14) suggested that β' i1 could rotate with respect to the rest of the β' clamp (β' 1-133 and β' 255-404) around a presumed hinge in-between these two domains. This, then, should allow opening of the DNA passageway toward the primary channel. Therefore, we analyzed the location of β' i1 with respect to the rest of the β' clamp in all variants of *M. smegmatis* RNAP complexes. The mutual orientations of these two domains stay rigid, and thus we conclude that there is no molecular hinge between these domains that would allow β' i1 any substantial movements with respect to the rest of the β' clamp. Nevertheless, the lack of any cryo-EM density for the β' il RNAP-distal part (the tip) then suggests that the tip may display some flexibility relative to the RNAP-proximal part of this domain. This flexibility may play a role in accommodation of DNA in the primary channel. Together, this confirms that solely the molecular hinge of the β' switch regions allows for flexible arrangement of the β' clamp and the adjacent β' il, and this opens the enzyme pincer enough to let double-stranded DNA pass through and accommodate it into the primary channel and the active site, similar to what has been recently reported for *E. coli* RNAP with a partially loaded DNA promoter (18, 19). It should be noted that in the case of the holoenzyme, part of the primary channel might be occupied by the sigma factor and thus reduce the actual space available for DNA. This might be a reason why RNAP molecules exhibit conformations where the pincer is splayed to around 30 Å.

(iii) Interaction of the RNAP core with σ^A and movement of $\sigma^A_{1,1}$. Our cryo-EM structure of the *M. smegmatis* holoenzyme lacks cryo-EM densities for the $\sigma^A_{3,2}$ and σ^A_4

TABLE 2 Positioning of the nonconserved domain β' i1 of *M. smegmatis* RNAP: distances between β Leu266C α and β' Arg214C α (*M. smegmatis*) and β Leu275C α and β' Arg214C α (*M. tuberculosis*)^a

RNAP complex	Organism	Distance (Å)	Residues	PDB entry	Source or reference
Core1	<i>M. smegmatis</i>	39.9	L266-R214	6F6W ^{EM}	This study
Core2	<i>M. smegmatis</i>	45.3	L266-R214	NA	This study
Holo1	<i>M. smegmatis</i>	38.3	L266-R214	6EYD ^{EM}	This study
Holo2	<i>M. smegmatis</i>	46.2	L266-R214	NA	This study
holo + RbpA	<i>M. tuberculosis</i>	38.7	L275-R214	6C05 ^{EM}	11
holo + LpM	<i>M. tuberculosis</i>	47.2	L275-R214	6FBV ^{EM}	23
holo + RbpA + Fdx	<i>M. tuberculosis</i>	45.3	L275-R214	6C06 ^{EM}	11
holo + RbpA + Fdx + up	<i>M. tuberculosis</i>	43.3	L275-R214	6BZO ^{EM}	11
holo + RbpA + up + down	<i>M. tuberculosis</i>	23.7	L275-R214	6C04 ^{EM}	11
holo + down	<i>M. tuberculosis</i>	21.9	L275-R214	5UHA ^{XR}	14
holo + down + Rif	<i>M. tuberculosis</i>	22	L275-R214	5UHB ^{XR}	14
holo + down + D-AAP1	<i>M. tuberculosis</i>	21.4	L275-R214	5UHE ^{XR}	14
holo + down + D-IX336	<i>M. tuberculosis</i>	21.7	L275-R214	5UHF ^{XR}	14
holo + down + Rif + D-AAP1	<i>M. tuberculosis</i>	21.2	L275-R214	5UHG ^{XR}	14
holo + RbpA + up	<i>M. smegmatis</i>	21.4	L266-R214	5V18 ^{XR}	12
RP _O + RbpA	<i>M. smegmatis</i>	21.8	L266-R214	5V15 ^{XR}	13
holo + down + 2 nt RNA	<i>M. tuberculosis</i>	22.2	L275-R214	5UH9 ^{XR}	14
holo + down + 3 nt RNA	<i>M. tuberculosis</i>	22.2	L275-R214	5UH5 ^{XR}	14
holo + down + 4 nt RNA	<i>M. tuberculosis</i>	22.1	L275-R214	5UH8 ^{XR}	14
holo + down + Rif + 2 nt RNA	<i>M. tuberculosis</i>	22.5	L275-R214	5UH6 ^{XR}	14
holo + down + Rif + 3 nt RNA	<i>M. tuberculosis</i>	22.1	L275-R214	5UHC ^{XR}	14
holo + down + Rif + 4 nt RNA	<i>M. tuberculosis</i>	21.9	L275-R214	5UHD ^{XR}	14

^a“up” stands for DNA upstream fork; “down” stands for DNA downstream fork. nt, nucleotide(s); Fdx, fidaxomicin; LpM, lipiarmycin A3; Rif, rifampin; D-AAP1, *N*- α -(benzenecarbonyl)-*N*-(2-methylphenyl)-*D*-phenylalaninamide; D-IX336, *N*-(2-methylphenyl)-*N*- α -(selenophene-2-carbonyl)-*D*-phenylalaninamide; NA, not applicable. Superscripts: XR, X-ray; EM, cryo-EM.

C-terminal regions (Fig. 3), which usually protrude through RNA exit channel in holoenzyme structures of bacterial polymerases (12). The constituents of the RNA exit channel cleft, the region of the β' flap, are generally not very well defined in the presented electron microscopic structures, and the β -flap tip (β 802 to 822) is missing entirely. The RNA exit cleft itself is wider compared to holoenzymes with a defined $\sigma_{3.2-4}$ region. This may be a consequence of the absence of other factors (like RpbA and/or antibacterial compounds) that were present in previous mycobacterial holoenzyme structures (11, 23).

A major difference exists in sequence and structure between *E. coli* $\sigma^{70}_{1.1}$ (including its preceding linker) and the equivalent N-terminal part of the σ^A region (σ^A_N) from mycobacteria, and this is reflected in interactions of this region with RNAP (Fig. 3). Whereas the *E. coli* $\sigma^{70}_{1.1}$ was identified in the primary channel in the RNAP holoenzyme (8), it still remained unclear whether mycobacterial σ^A_N was directed outside by its $\sigma^A_{N\text{-helix}}$ (12, 14).

In previously published mycobacterial RNAP closed conformation structures (i.e., PDB entries 5V18 [12]; 5V15 [13]; 5UHA, 5UHB, 5UHE, 5UHF, 5UHG, 5UH9, 5UH5, 5UH8, 5UH6, 5UHC, and 5UHD [14]; and 6C04 [11]), the σ^A region 1.1 appeared to be mostly unstructured with only the $\sigma^A_{N\text{-helix}}$ (σ^A 149 to 162) well defined and positioned outside the RNAP active site cleft on periphery in between the β and β' subunit claws, parallel to the taxon-specific β' i1 finger (Fig. 3, middle). Two different proposals have been made regarding the position of mycobacterial σ^A_N in RNAP holoenzyme: first, it was proposed that, that unlike *E. coli* $\sigma^{70}_{1.1}$, *M. smegmatis* σ^A_N may never reside in the RNAP active-site cleft (12), whereas it was previously proposed that, like Eco *E. coli* $\sigma^{70}_{1.1}$, mycobacterial σ^A_N resides in the active-site cleft in RNAP holoenzyme (14).

Consistent with the second model (14), the *M. tuberculosis* $\sigma^A_{N\text{-helix}}$ (defined by residues σ^A 207 to 225 that is preceded by a short defined loop [σ^A 202 to 206], which is buttressed by β domain 2) was directed into the RNAP active-site cleft in several recent open-clamp conformation structures of *M. tuberculosis* RNAP (PDB entries 6C05, 6C06, and 6BZO [11] and in cryo-EM density EMDDataBank [EMD] entry EMD-4230 [23]). However, this interesting fact was not mentioned (11, 23).

FIG 3 The σ^A N-helix changes position within the RNAP. (Top left) The N-terminal α -helix (red) of σ^A (amino acids [aa] 149 to 162) is located within the primary channel of the *M. smegmatis* open-conformation holoenzyme RNAP. The σ^A N-helix is positioned between the β' coiled-coil and the inner edge of β domain 2 (Table S2) and oriented perpendicular to the nonconserved domain (NCD) β' i1 finger. (Top middle) The N-terminal α -helix (red) of σ^A (aa 149 to 162) in the closed-conformation mycobacterial RNAP structures (12) is positioned on the periphery in between the β and σ' subunit claws, parallel to the NCD β' i1 finger. (Top right) $\sigma 701.1$ from *E. coli* in complex with RNAP positioned in the primary channel of RNAP (8). RNAP subunits are colored as follows: β , yellow; β' , green; α , gray; α' , cyan; ω , orange; $\sigma^A/70$, magenta; the N-terminal part of $\sigma^A/70$, dark blue; σ^A N-helix, red. (Bottom row) Detailed views of the respective $\sigma^A/70$ N helix regions from the structures in the upper row.

We here describe here that in the open-clamp holoenzyme structure the $\sigma^A_{N\text{-helix}}$ is present within the primary channel (Fig. 3, left; see Fig. S7). There, the $\sigma^A_{N\text{-helix}}$ (namely, amino acids Ala149-Ala161) is positioned between the β and β' subunit claws, spanning from the β' coiled-coil to the inner edge of β domain 2 (Table S2) and perpendicular to β' i1. The $\sigma^A_{N\text{-helix}}$ must be displaced from this position in between the RNAP clamp to permit a conformational change during RP_o formation, since it would clash with incoming or even already loaded DNA. This is in agreement with the observation that deletion of the 1.1 region influences the stability of mycobacterial RP_o (14). The rest of the 1.1 region is not defined, but it is clear the *M. smegmatis* $\sigma^A_{N\text{-helix}}$ plays a pivotal role in positioning of the N-terminal part of the σ^A protein and, in the case of the holoenzyme the $\sigma^A_{N\text{-helix}}$ probably directs at least a part of it into the downstream region of the primary channel.

The positioning and movement of region 1.1 must also depend on the net negative charge within the σ^A_N disordered part (12) that likely drives its positioning in between the positively charged inner surface of the clamp and β' i1. On the other hand, the positively charged first part of σ^A_N might interact with the negatively charged surface of RNAP and the outer side of β' i1. A potential functional role of this directional charge arrangement needs to be further clarified.

To conclude, the mycobacterial RNAP core and holoenzyme structures presented here, combined with other reported structures of mycobacterial RNAP, complete the overall picture of the main conformational forms of this enzyme.

MATERIALS AND METHODS

Bacterial strains and plasmids. Vector pAC22 (a pET28a derivative, a gift from Bob Landick Lab) containing *Mycobacterium bovis* RNAP core subunit genes was used as a template to create an analogous

vector containing *M. smegmatis* RNAP core genes. *rpoA*, *rpoZ*, *rpoB*, and *rpoC* were amplified by PCR using primers listed in Table S3 in the supplemental material and genomic DNA of *Mycobacterium smegmatis* mc²155 as a template. The *M. smegmatis* *rpoA* gene was inserted to the pAC22 vector via XbaI and PacI restriction sites. The *rpoC* gene was assembled from two amplified fragments via the KpnI restriction site, and the complete gene, containing an 8×His tag on the 3' end, was then inserted into the pAC22 vector via BamHI and Ascl restriction sites. The *rpoB* gene was inserted together with a 9-amino-acid polylinker via NotI and Ascl restriction sites and connected in frame to *rpoC*. Finally, the *rpoZ* gene was inserted into the pAC22 vector via the PacI and NotI restriction sites. The final vector, pRMS4, encodes a polycistronic transcript for expression of all five RNAP core subunits (Fig. S1A). It contains a β' C-terminal 8×His tag for affinity purification. Expression is driven from a T7 RNAP-dependent promoter.

pSigA-His6 encoding *M. smegmatis* Sigma A factor was prepared as follows. The *mysA* gene was cloned into the pET22b+ expression vector via EcoRI and XhoI restriction sites. The *mysA* gene thus obtained a 6×His tag at the 3' end.

***M. smegmatis* RNAP core purification.** *E. coli* strain BL21(DE3) was transformed with the pRMS4(kanR) plasmid. To express and purify *M. smegmatis* RNAP, cultures were incubated at 37°C until reaching an optical density at 600 nm (OD₆₀₀) of ~0.8; expression was induced with 500 μM IPTG (isopropyl-β-D-thiogalactopyranoside) at 17°C for 16 h. Cells were lysed using sonication by Sonic dismembrator model 705 (Fisher Scientific) in a lysis buffer containing 50 mM NaH₂PO₄-Na₂HPO₄ (pH 8; 4°C), 300 mM NaCl, 2.5 mM MgCl₂, 30 mM imidazole, 5 mM β-mercaptoethanol, EDTA-free protease inhibitor cocktail (Roche), RNase A (Sigma), DNase I (Sigma), and lysozyme (Sigma). Clarified lysate was loaded onto a HisTrap FF crude column (GE Healthcare), and proteins were eluted with a linear gradient of imidazole to a final concentration of 400 mM over 20 column volumes. The *M. smegmatis* RNAP core elution fractions were pooled and dialyzed to 20 mM Tris-HCl (pH 8; 4°C), 1 M NaCl, 5% (vol/vol) glycerol, and 4 mM dithiothreitol (DTT) for 20 h. The protein was further purified on a XK 26/70 Superose 6-pg column (GE Healthcare) equilibrated in 20 mM Tris-HCl (pH 8; 4°C), 300 mM NaCl, 5% (vol/vol) glycerol, and 4 mM DTT. The final *M. smegmatis* RNAP core fractions were eluted at 6 μM, aliquoted, flash-frozen in liquid nitrogen, and then stored at -80°C.

SigA (σ^A) factor purification. *E. coli* strain BL21(DE3) was transformed with pET22b+(ampR) plasmid derivative encoding the *M. smegmatis* σ^A factor fusion with a C-terminal 8×His tag, transcribed from a T7 promoter. To express and purify σ^A, the culture was incubated at 37°C until the OD₆₀₀ reached ~0.8; expression was induced with 500 μM IPTG at 17°C for 16 h. Cells were lysed using sonication by Sonic dismembrator model 705 (Fisher Scientific) in a lysis buffer containing 50 mM NaH₂PO₄-Na₂HPO₄ (pH 8; 4°C), 300 mM NaCl, 2.5 mM MgCl₂, 30 mM imidazole, 5 mM β-mercaptoethanol, EDTA-free protease inhibitor cocktail (Roche), RNase A (Sigma), DNase I (Sigma), and lysozyme (Sigma). Clarified lysate was loaded onto a HisTrap FF crude column (GE Healthcare), and proteins were eluted with a linear gradient of imidazole to the final concentration of 400 mM over 20 column volumes. Fractions containing σ^A were pooled and dialyzed for 20 h against the dialysis buffer containing 20 mM Tris-HCl (pH 8; 4°C), 200 mM NaCl, 5% (vol/vol) glycerol, and 4 mM DTT and then loaded onto a HiTrap heparin HP column (GE Healthcare), followed by column equilibration in dialysis buffer containing 350 mM NaCl. The protein was then eluted in one step with dialysis buffer containing 450 mM NaCl. Finally, the elution fractions were directly subjected to a size exclusion chromatography using a Superdex 75 column (GE Healthcare) equilibrated in 20 mM Tris-HCl (pH 8; 4°C), 300 mM NaCl, 5% (vol/vol) glycerol, and 4 mM DTT. The σ^A protein was eluted at 66 μM, aliquoted, flash-frozen in liquid nitrogen, and then stored at -80°C.

***In vitro* transcription assay.** *In vitro* transcriptions were performed with two σ^A-dependent promoters: *Pveg*, a *B. subtilis* vegetative promoter (28), and *PrnAPCL1*, a *M. smegmatis* ribosomal promoter (29). Multiple-round transcription assays were carried out as described previously (29). Briefly, the RNAP core was reconstituted with σ^A for 15 min at 37°C in 1× standard transcription buffer [STB; 5 mM Tris-HCl (pH 8.0), 1 mM Mg(C₂H₃O₂)₂, 20 μM DTT, 10 mM KCl, and 50 μg/ml bovine serum albumin] to form holoenzyme at a ratio of 1:5 (RNAP to σ^A). Subsequently, a mixture of promoter (final concentration, 50 nM) and NTPs (the ATP, CTP, and GTP concentrations were 200 μM; the UTP concentration was 10 μM plus 2 μM radiolabeled [α-³²P]UTP) in STB was prepared. Finally, transcription reactions were initiated with holoenzyme (final concentration, 200 nM). All transcriptions were allowed to proceed for 15 min at 37°C. After 15 min, transcription was stopped by the addition of 1× stop solution (95% formamide, 20 mM EDTA [pH 8.0], 0.05% bromophenol blue, 0.05% xylene cyanol). Samples were loaded onto a 7% polyacrylamide gel and electrophoresed. The gels were scanned by using a Molecular Imager_FX (Bio-Rad). Data were analyzed with QuantityOne software (Bio-Rad).

RNAP core and σ^A factor *in vitro* reconstitution for cryo-EM. To assemble the holoenzyme consisting of the RNAP core and σ^A, the individual proteins were mixed at a molar ratio of 1:3, respectively. The *in vitro* reconstitutions were carried out at 4°C, and the reconstitution mixture was incubated for 15 min. Then, 50 μl of the reconstitution mixture was injected onto a Superose 6 Increase 3.2/300 column (GE Healthcare) equilibrated in 20 mM Tris-HCl (pH 7.8; 4°C), 150 mM NaCl, 10 mM MgCl₂, and 2 mM DTT. Next, 50-μl fractions were collected, and the protein was eluted at 1.5 μM (~17-fold above the estimated K_d). The RNAP core and RNAP σ^A comigrated in size exclusion chromatography (SEC) in a single peak (Fig. S1B).

Electron microscopy. Reconstituted complexes were processed immediately after the last size exclusion chromatography step. Complexes were diluted to ~110 nM, and aliquots of 3 μl were applied to Quantifoil R1.2/1.3 Au 300 mesh grids coated with graphene oxide (Graphene Supermarket). The grids were incubated for 30 s at 4°C and 100% humidity, blotted for 2 s, and then plunged into liquid ethane using an FEI Vitrobot IV. The grids were loaded into an FEI Tecnai Polara electron microscope operated at an accelerating voltage of 300 kV and equipped with a Falcon III direct detector. Micrographs were

taken using EPU software (FEI) at a nominal magnification of $\times 78,000$, yielding a pixel size of 1.34 Å per pixel at the specimen level.

Each image was exposed for 1.8 or 2 s at dose rate of around 36 or 26 electrons/Å²/s, respectively. The beam used was larger (ca. 1.6 μm in diameter) than the Quantifoil hole, illuminating the carbon all around the hole. Defocus range was set at -1.8 to -3.8 μm. An in-house built system was used to intercept videos from the detector, and 56 (1.8-s exposure) or 62 (2-s exposure) movie frames were recorded for each image (30).

Image processing. Movie frames were first aligned by the MotionCor2 (31) program before subsequent processing. First three movie frames of total dose of approximately 3 electrons/Å² were discarded, mimicking a preexposure. The rest of the frames were aligned using a dose-weighting scheme of MotionCor2 program. The total dose used was 52 electrons/Å² (2-s exposure) or 64 electrons/Å² (1.8-s exposure), which equals around 1 electron/Å² in one movie frame. Contrast transfer function parameters were calculated with Gctf (32) from the motion-corrected and not dose-weighted images. Initially, particles (in 208×208 pixels box size) were selected automatically by Gautomatch (provided in-house by Kai Zhang [<http://www.mrc-lmb.cam.ac.uk/kzhang>]) from a small portion of the data set (~200 micrographs [see Fig. S2A]). The relatively high dose enabled sufficient contrast to allow particle picking even in low defocused micrographs. After 2D classification using RELION 2.0 (30), eight images in different views were selected from the 2D averages and used as reference for automatic particle picking for the whole data set by RELION 2.0 or Gautomatch.

The following procedures were performed in RELION to exclude bad particles from the final reconstruction. First, particles in each micrograph were displayed and manually screened to delete wrongly picked areas. Second, 2D classification was performed iteratively to obtain the accuracy of the rotations of most of the 2D classes of $\sim 2.6^\circ$. Particles in bad classes with poor structural features were removed.

The resultant particles (~540 thousands) were then used for 3D classifications. The first round of 3D classification restricted to eight classes was performed using X-ray crystal structure of *E. coli* RNA polymerase σ^{70} holoenzyme (PDB entry 4YG2) as a 60-Å low-pass filtered initial model. Classification was done during three rounds of 25 iterations each, using the regularization parameter $T=4$ and restricted to 12 classes. During the second and third rounds, local angular searches were performed to 3.5° and 1.8° to clearly separate structural species. Most abundant classes were pooled, and a second round of classification was performed, again restricted to eight classes. The final accuracy of the rotational alignment was $\sim 2^\circ$, and the translational alignment accuracy was 0.7 pixels. Distinguishable species of RNAP core and σ^A -bound holoenzyme were identified and further classified separately. A summary of particles used for final refinements is listed in Table S1. We used statistical movie processing in RELION (30) to estimate the per-particle beam induced movement for all movie frames with running averages of five movie frames and a standard deviation of 1 pixel for the translational alignments. For each class independently, we used a particle-polishing step (30), which includes B-factor weighting to estimate the dose- and resolution-dependent radiation damage for each movie frame. In this step, the signal from the very late frames, with very high radiation damage, was damped. The final particle sets were 3D autorefined and postprocessed in RELION. The local resolution (Fig. S3) was calculated using the RELION software package. The angular distribution of particles was calculated using the cryoEF software package (24).

Model building and refinement. Atomic models of protein parts of *M. smegmatis* RNAP species (Fig. 1 and 2) were generated according to the recent crystal structure of *M. smegmatis* RP_O and TIC (PDB entries 5V15 and 5V18). Protein chains of the whole complexes were first rigid-body fit into cryo-EM density by Molrep (33). Individual subdomains of *M. smegmatis* RNAP were then assigned according to the rigid modules previously described in *T. thermophilus* (26) (Table S2) and fit into the cryo-EM density by JiggleFit tool (34) in Coot (35). Best fits of the subdomains from both 5V15 and 5V18 were chosen according to a correlation coefficient in the JiggleFit tool. The atomic models building for the second most abundant conformation forms of RNAP core and holoenzyme were stopped at this point.

The models were then iteratively improved by manual building in Coot and refinement in Refmac (36). For the Refmac refinement, the model atomic distance were restrained according to the crystal structure of *M. smegmatis* RP_O (5V15) in ProSMART (37). The criterion for refinement completion was a stabilized R factor and low RMS of bond angles and lengths (Table S1).

Accession number(s). The EM maps and atomic coordinates of *M. smegmatis* RNA polymerase core (PDB entry 6F6W and EMD entry EMD-4192) and holoenzyme (PDB entry 6EYD and EMD entry EMD-3983) have been deposited in the Protein Data Bank (<http://www.pdb.org>) and the EMDataBank (<http://emdatabank.org/>).

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at <https://doi.org/10.1128/JB.00583-18>.

SUPPLEMENTAL FILE 1, PDF file, 1.4 MB.

ACKNOWLEDGMENTS

We thank Alan Roy Fersht for material and scientific support of this work in his laboratory. We also thank Xiaochen Bai and Kai Zhang for advice with cryo-EM data processing and Christos Savva and Greg McMullan for help with the data collection. We

thank Jake Grimmett and Toby Darling for assistance with the Laboratory of Molecular Biology, MRC, computing cluster.

This study was funded by the Czech Science Foundation grants 17-03419S (to L.K.) and 13-27150P (to J.H.) and by grant GAUK 794317 (to J.P.).

We declare that we have no conflicts of interest with the contents of this article.

T.K. expressed and purified proteins, reconstituted the RNAP complexes, prepared cryo-EM grids, collected EM data, performed image processing and reconstruction, built models, made figures, and wrote the manuscript. J.P. and J.H. cloned the expression vectors, expressed and purified proteins, and performed the transcription assays. J.H. and H.Š. developed expression and purification protocols. I.B. prepared the figures and wrote the manuscript. L.K. conceived the project, organized the collaboration, and wrote the manuscript.

REFERENCES

1. Molodtsov V, Scharf NT, Stefan MA, Garcia GA, Murakami KS. 2017. Structural basis for rifamycin resistance of bacterial RNA polymerase by the three most clinically important RpoB mutations found in *Mycobacterium tuberculosis*. *Mol Microbiol* 103:1034–1045. <https://doi.org/10.1111/mmi.13606>.
2. Ma C, Yang X, Lewis PJ. 2016. Bacterial transcription as a target for antibacterial drug development. *Microbiol Mol Biol Rev* 80:139–160. <https://doi.org/10.1128/MMBR.00055-15>.
3. Sutherland C, Murakami KS. 2018. An introduction to the structure and function of the catalytic core enzyme of *Escherichia coli* RNA polymerase. *EcoSal Plus* 8. <https://doi.org/10.1128/ecosalplus.ESP-0004-2018>.
4. Paget MS. 2015. Bacterial sigma factors and anti-sigma factors: structure, function and distribution. *Biomolecules* 5:1245–1265. <https://doi.org/10.3390/biom5031245>.
5. Bae B, Chen J, Davis E, Leon K, Darst SA, Campbell EA. 2015. CarD uses a minor groove wedge mechanism to stabilize the RNA polymerase open promoter complex. *Elife* 4:e08505. <https://doi.org/10.7554/eLife.08505.001>.
6. Hu Y, Morichaud Z, Perumal AS, Roquet-Baneres F, Brodolin K. 2014. *Mycobacterium* RbpA cooperates with the stress-response sigmaB subunit of RNA polymerase in promoter DNA unwinding. *Nucleic Acids Res* 42:10399–10408. <https://doi.org/10.1093/nar/gku742>.
7. Sudalaiyadum Perumal A, Vishwakarma RK, Hu Y, Morichaud Z, Brodolin K. 2018. RbpA relaxes promoter selectivity of *Mycobacterium tuberculosis* RNA polymerase. *Nucleic Acids Res* 46:10106–10118. <https://doi.org/10.1093/nar/gky714>.
8. Bae B, Davis E, Brown D, Campbell EA, Wigneshweraraj S, Darst SA. 2013. Phage T7 Gp2 inhibition of *Escherichia coli* RNA polymerase involves misappropriation of sigma70 domain 1.1. *Proc Natl Acad Sci U S A* 110:19772–19777. <https://doi.org/10.1073/pnas.1314576110>.
9. Schwartz EC, Shekhtman A, Dutta K, Pratt MR, Cowburn D, Darst S, Muir TW. 2008. A full-length group 1 bacterial sigma factor adopts a compact structure incompatible with DNA binding. *Chem Biol* 15:1091–1103. <https://doi.org/10.1016/j.chembiol.2008.09.008>.
10. Zachrdla M, Padrt P, Rabatinova A, Sanderova H, Barvik I, Krasny L, Zidek L. 2017. 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. <https://doi.org/10.1074/jbc.M117.784074>.
11. Boyaci H, Chen J, Lilic M, Palka M, Mooney RA, Landick R, Darst SA, Campbell EA. 2018. Fidaxomicin jams *Mycobacterium tuberculosis* RNA polymerase motions needed for initiation via RbpA contacts. *Elife* 7:e34823. <https://doi.org/10.7554/eLife.34823>.
12. Hubin EA, Fay A, Xu C, Bean JM, Saecker RM, Glickman MS, Darst SA, Campbell EA. 2017. Structure and function of the mycobacterial transcription initiation complex with the essential regulator RbpA. *Elife* 6:e22520. <https://doi.org/10.7554/eLife.22520>.
13. Hubin EA, Lilic M, Darst SA, Campbell EA. 2017. Structural insights into the mycobacteria transcription initiation complex from analysis of X-ray crystal structures. *Nat Commun* 8:16072. <https://doi.org/10.1038/ncomms16072>.
14. Lin W, Mandal S, Degen D, Liu Y, Ebright YW, Li S, Feng Y, Zhang Y, Mandal S, Jiang Y, Liu S, Gigliotti M, Talaue M, Connell N, Das K, Arnold E, Ebright RH. 2017. Structural basis of *Mycobacterium tuberculosis* transcription and transcription inhibition. *Mol Cell* 66:169–179. <https://doi.org/10.1016/j.molcel.2017.03.001>.
15. Chakraborty A, Wang D, Ebright YW, Korlann Y, Kortkhonjia E, Kim T, Chowdhury S, Wigneshweraraj S, Irschik H, Jansen R, Nixon BT, Knight J, Weiss S, Ebright RH. 2012. Opening and closing of the bacterial RNA polymerase clamp. *Science* 337:591–595. <https://doi.org/10.1126/science.1218716>.
16. Darst SA, Opalka N, Chacon P, Polyakov A, Richter C, Zhang G, Wriggers W. 2002. Conformational flexibility of bacterial RNA polymerase. *Proc Natl Acad Sci U S A* 99:4296–4301. <https://doi.org/10.1073/pnas.052054099>.
17. Ruff EF, Drennan AC, Capp MW, Poulos MA, Artsimovitch I, Record MT, Jr. 2015. *E. coli* RNA polymerase determinants of open complex lifetime and structure. *J Mol Biol* 427:2435–2450. <https://doi.org/10.1016/j.jmb.2015.05.024>.
18. Glyde R, Ye F, Jovanovic M, Kotta-Loizou I, Buck M, Zhang X. 2018. Structures of Bacterial RNA polymerase complexes reveal the mechanism of DNA loading and transcription initiation. *Mol Cell* 70:1111–1120. <https://doi.org/10.1016/j.molcel.2018.05.021>.
19. Glyde R, Ye F, Darbari VC, Zhang N, Buck M, Zhang X. 2017. Structures of RNA polymerase closed and intermediate complexes reveal mechanisms of DNA opening and transcription initiation. *Mol Cell* 67:106–116. <https://doi.org/10.1016/j.molcel.2017.05.010>.
20. Browning DF, Busby SJ. 2016. Local and global regulation of transcription initiation in bacteria. *Nat Rev Microbiol* 14:638–650. <https://doi.org/10.1038/nrmicro.2016.103>.
21. Krasny L, Tiserova H, Jonak J, Rejman D, Sanderova H. 2008. The identity of the transcription +1 position is crucial for changes in gene expression in response to amino acid starvation in *Bacillus subtilis*. *Mol Microbiol* 69:42–54. <https://doi.org/10.1111/j.1365-2958.2008.06256.x>.
22. Lane WJ, Darst SA. 2010. Molecular evolution of multisubunit RNA polymerases: sequence analysis. *J Mol Biol* 395:671–685. <https://doi.org/10.1016/j.jmb.2009.10.062>.
23. Lin W, Das K, Degen D, Mazumder A, Duchi D, Wang D, Ebright YW, Ebright RY, Sineva E, Gigliotti M, Srivastava A, Mandal S, Jiang Y, Liu Y, Yin R, Zhang Z, Eng ET, Thomas D, Donadio S, Zhang H, Zhang C, Kapanidis AN, Ebright RH. 2018. Structural basis of transcription inhibition by fidaxomicin (lipiarmycin A3). *Mol Cell* 70:60–71. <https://doi.org/10.1016/j.molcel.2018.02.026>.
24. Naydenova K, Russo CJ. 2017. Measuring the effects of particle orientation to improve the efficiency of electron cryomicroscopy. *Nat Commun* 8:629. <https://doi.org/10.1038/s41467-017-00782-3>.
25. Vassylyev DG, Vassylyeva MN, Perederina A, Tahirov TH, Artsimovitch I. 2007. Structural basis for transcription elongation by bacterial RNA polymerase. *Nature* 448:157–162. <https://doi.org/10.1038/nature05932>.
26. Tagami S, Sekine S, Kumarevel T, Hino N, Murayama Y, Kamegamori S, Yamamoto M, Sakamoto K, Yokoyama S. 2010. Crystal structure of bacterial RNA polymerase bound with a transcription inhibitor protein. *Nature* 468:978–982. <https://doi.org/10.1038/nature09573>.
27. Duchi D, Mazumder A, Malinen AM, Ebright RH, Kapanidis AN. 2018. The RNA polymerase clamp interconverts dynamically among three states and is stabilized in a partly closed state by ppGpp. *Nucleic Acids Res* 46:7284–7295. <https://doi.org/10.1093/nar/gky482>.
28. Sojka L, Kouba T, Barvik I, Sanderova H, Maderova Z, Jonak J, Krasny

- L. 2011. Rapid changes in gene expression: DNA determinants of promoter regulation by the concentration of the transcription initiating NTP in *Bacillus subtilis*. *Nucleic Acids Res* 39:4598–4611. <https://doi.org/10.1093/nar/gkr032>.
29. China A, Tare P, Nagaraja V. 2010. Comparison of promoter-specific events during transcription initiation in mycobacteria. *Microbiology* 156: 1942–1952. <https://doi.org/10.1099/mic.0.038620-0>.
30. Scheres SH. 2012. A Bayesian view on cryo-EM structure determination. *J Mol Biol* 415:406–418. <https://doi.org/10.1016/j.jmb.2011.11.010>.
31. Zheng SQ, Palovcak E, Armache JP, Verba KA, Cheng Y, Agard DA. 2017. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat Methods* 14:331–332. <https://doi.org/10.1038/nmeth.4193>.
32. Zhang K. 2016. Gctf: real-time CTF determination and correction. *J Struct Biol* 193:1–12. <https://doi.org/10.1016/j.jsb.2015.11.003>.
33. Vagin A, Teplyakov A. 2010. Molecular replacement with MOLREP. *Acta Crystallogr D Biol Crystallogr* 66:22–25. <https://doi.org/10.1107/S0907444909042589>.
34. Brown A, Long F, Nicholls RA, Toots J, Emsley P, Murshudov G. 2015. Tools for macromolecular model building and refinement into electron cryo-microscopy reconstructions. *Acta Crystallogr D Biol Crystallogr* 71: 136–153. <https://doi.org/10.1107/S1399004714021683>.
35. Emsley P, Lohkamp B, Scott WG, Cowtan K. 2010. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66:486–501. <https://doi.org/10.1107/S0907444910007493>.
36. Murshudov GN, Skubak P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, Winn MD, Long F, Vagin AA. 2011. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr* 67:355–367. <https://doi.org/10.1107/S0907444911001314>.
37. Nicholls RA, Long F, Murshudov GN. 2012. Low-resolution refinement tools in REFMAC5. *Acta Crystallogr D Biol Crystallogr* 68:404–417. <https://doi.org/10.1107/S090744491105606X>.