Structural basis of ribosomal RNA transcription regulation

Yeonoh Shin[1,a], M. Zuhaib Qayyum[1,a], Danil Pupov[2], Daria Esyunina[2], Andrey Kulbachinskiy[2] and Katsuhiko S. Murakami[1,*]

[1]Department of Biochemistry and Molecular Biology, Pennsylvania State University, University Park, PA 16802

[2]Institute of Molecular Genetics, Russian Academy of Sciences, Moscow 123182, Russia

[a]Y.S. and M.Z.Q. contributed equally to this work.

[*]Correspondence: kum14@psu.edu

Ribosomal RNA (rRNA) is the most highly expressed gene in rapidly growing bacteria and is drastically downregulated under stress conditions by the global transcriptional regulator DksA and the alarmone ppGpp. To reveal the mechanism of highly regulated rRNA transcription, we determined cryo-electron microscopy structures of the *Escherichia coli* RNA polymerase (RNAP) σ^{70} holoenzyme at different steps of rRNA promoter recognition with and without DksA/ppGpp. RNAP contacts the UP element of rRNA promoter using the dimerized α subunit carboxyl-terminal domain and scrunches the template DNA with the σfinger and β'lid to select a transcription start site favorable for rRNA expression. Promoter DNA binding to RNAP induces conformational change of the σ domain 2 that opens a gate for DNA loading and ejects $\sigma_{1.1}$ from the RNAP cleft to facilitate open complex formation. DksA/ppGpp binding to RNAP also opens the DNA loading gate, but it is not coupled to $\sigma_{1.1}$ ejection and impedes the open complex formation of the rRNA promoter due to its G+C rich discriminator sequence. Mutations in $\sigma_{1.1}$ or the β'lid stabilize the RNAP and rRNA promoter complex and decrease its sensitivity to DksA/ppGpp. These results provide a molecular basis for exceptionally active rRNA transcription and for its vulnerability to DksA/ppGpp.

Bacteria sense the availability of nutrition and adjust ribosome biogenesis to optimize their growth. The rate of ribosome biogenesis is primarily determined by rRNA transcription [1,2], which constitutes as much as 70 % of total transcription and is initiated approximately every second from each of the seven operons (*rrn*A-E and *rrn*G-H) in *E. coli* during exponential growth [3]. However, it is drastically repressed under stress conditions such as nutrient-starved stationary phase [4]. rRNA expression is regulated at the initiation stage of RNA synthesis, including RNAP binding to promoter DNA, unwinding the DNA and escaping from the promoter.

The promoters (e.g., *rrnB*P1) for expressing rRNA operons are unique compared with other promoters, including 1) the A+T rich UP element located upstream of the -35 element (from -60 to -40); 2) the G+C rich discriminator sequence downstream of the -10 element (from -8 to -1); and 3) the transcription start site (TSS) located 9 bases downstream from the -10 element (**Fig. 1A, Supplemental Fig. 1A**). The UP element is recognized by the carboxyl-terminal domain of the α subunit (αCTD) and enhances rRNA transcription by more than 30-fold [5]. The G+C rich discriminator and unusual TSS selection make the open complex (RPo) unstable, but this facilitates

promoter escape by reducing the abortive RNA cycle prior to the productive RNA elongation stage [6]. These promoter elements play key roles in the wide range of rRNA transcription regulation between optimized and nonoptimized growth conditions.

rRNA transcription activity is regulated by the concentrations of two molecules - the initiating ribonucleotide (iNTP) (ATP in the case of *rrnB*P1) [7] and the bacterial alarmone ppGpp (guanosine tetraphosphate, aka "magic spot"), which is an allosteric effector of the RNAP-binding regulator DksA [8-10]. The presence of high iNTP allows RNAP to break contacts with the promoter in the intrinsically unstable rRNA RPo, allowing escape to transcription elongation. However, the iNTP-limited condition shifts the equilibrium to favor early intermediates, including the closed complex (RPc), which is further enhanced by DksA/ppGpp binding to RNAP [4]. The ppGpp concentration is increased under stress conditions, which enhances DksA-mediated rRNA transcription repression by stabilizing DksA in a functionally important binding mode.

The majority of bacterial RNAP-DNA complex structures determined by X-ray crystallography contain short promoter DNA fragments with a premelted transcription bubble that mimics RPo to maximize its stability [11,12]. These studies explained the structural basis of promoter recognition and transcript initiation but left unexplored the interactions of RNAP with duplex DNA around the UP element (via αCTDs) and the contacts with the -10 element (via σ domain 2, residues 96-127 and 373-456 in σ^{70}) in the RPc and the scrunched DNA bubble in the stressed RPo formed with the rRNA promoters.

Cryo-electron microscopy (cryo-EM) structures of the *E. coli* RNAP-*rpsT*P2 promoter complex with a DksA homolog TraR revealed the RPo formation pathway in the presence of TraR, with stepwise RNAP conformational changes [13]. However, the *rpsT*P2 promoter for expressing ribosomal protein S20 is distinct from the *rrnB*P1 promoter in that it contains a G+C rich UP element and the TSS 7 bases downstream from the -10 element; therefore, it could not reveal the pathway for rRNA promoter complex formation and the mechanism of rRNA transcription regulation. Here, we used cryo-EM to visualize the RNAP and *rrnB*P1 complex in the RPc and RPo stages and two intermediates with DksA/ppGpp on the way to RPo formation.

Cryo-EM structure of the RNAP and *rrnB*P1 promoter closed complex (RPc)

We preincubated RNAP with *rrnB*P1 promoter DNA (**Fig. 1A**) at 37 °C for 5 min prior to cryo-EM grid preparation. In a separate cryo-EM grid preparation, we also tested adding iNTPs (ATP

and the nonhydrolyzable nucleotide CMPCPP) to stabilize the RNAP-*rrnB*P1 complex. In the course of cryo-EM data processing, 3D classification revealed 2 distinct structures based on the differences in the UP element (from -60 to -40), the downstream DNA (from -14 and +20) and the conformation of the σ factor (Methods, **SFigs. 2 and 3**), corresponding to the RPc and RPo.

The first class represents the RNAP-*rrnB*P1 closed complex (RPc) with an overall resolution of 4.14 Å (**STable 1**). The cryo-EM density shows that RNAP binds the duplex DNA from -60 to +3 (**Fig. 1B, SFig. 4, SMovie 1**), but the density of downstream DNA beyond position +4 is not traceable. Instead, a second DNA binds to the RNAP cleft due to the ejection of $\sigma_{1.1}$ from the RNAP cleft during RPc formation as described below.

The cryo-EM density for both αCTDs (residues 248-329), the linkers (residues 236-247) connecting to αNTDs (residues 1-235), and the UP element DNA were traceable in the RPc, allowing us to investigate how each αCTD binds to the UP element unambiguously (**Fig. 1C, SMovie 1**). Two αCTDs form a head-to-tail dimer and bind DNA side-by-side in the middle of the UP element (-51 to -48 on nontemplate DNA (ntDNA) and -54 to -50 on template DNA (tDNA)), which is in good agreement with the DNA footprinting results [14]. Although α subunits form a homodimer, two α subunits play different roles in RNAP, with one (α_1) adjacent to the β subunit and the other (α_2) adjacent to the β'. Compared to the α_2CTD, the α_1CTD is positioned proximally to the -35 element, which is consistent with the DNA cleavage by hydroxyl radicals from chelated Fe at each of the two αCTDs [15]. The side chains of R265 and N294 from both αCTDs are inserted into the DNA minor groove, and basic residues (K291 and K298) are involved in salt bridges with the DNA phosphate backbone (**SFig. 5**). The linkers of both α subunits are fully extended, and slight DNA bending centered at the -37 position is required for the αCTDs binding to the UP element (**Fig. 1C**). Consistent with this observation, shortening of the linkers by only three amino acids reduces *rrnB*P1 transcription [16]. Several studies have proposed that distant upstream DNA (near the -100 position) warps around RNAP on the RPo formation pathway, and the interaction of αCTDs and the UP element is one of the major driving forces for this DNA wrapping [17,18]. However, αCTDs do not bend the DNA around its binding site in the RPc structure, indicating that these contacts may not contribute to the DNA wrapping by RNAP.

The RPc structure shows how σ_2 binds the duplex form of the -10 element. The DNA encoding the -10 element is anchored by σ domain 2 and slightly bends around the upstream edge of the -10

element, allowing the downstream part beyond the -10 element to reach the other side of the RNAP cleft comprising the β protrusion domain (**Fig. 1B, SMovie 1**). The σ region 2.3 ($\sigma_{2.3}$, residues 417-434) contacts the -10 element by fitting into the DNA major groove without any amino acid–DNA base interaction, indicating that $\sigma_{2.3}$ recognizes the shape and/or curvature around the -10 element. This finding is in agreement with the previous proposal [19] that σ^{70} does not contact the -10 element DNA bases when it is in duplex form.

Cryo-EM structure of the RNAP and *rrnB*P1 promoter open complex (RPo)

The RNAP-*rrnB*P1 open complex (RPo) was determined with an overall resolution of 3.5 Å (**STable 1**). The cryo-EM density covers DNA from -44 to +20, including an open bubble from -13 to +2 and the downstream DNA accommodated in the RNAP cleft (**Fig. 2A, SFig. 4, SMovie 2**). In contrast to the RPc, αCTDs and the UP element are disordered. Basic residues in the βlobe (K163, K169, K191, R202 and K203), β'jaw (K1151, K1167, K1170, K1172 and R1174) and β'clamp (R133, K213, K216, K219 and R311) participate in the interaction with downstream DNA to stabilize the RPo (**Fig. 2B**). The importance of these interactions in rRNA transcription regulation is supported by the isolation of Δ*dksA* suppressor mutations in these domains [20,21]. The βgate loop (βGL, residues 368-378) in the βlobe domain contacts σ regions 1.1 ($\sigma_{1.1}$, residues 1-95) and 1.2 ($\sigma_{1.2}$, residues 96-127) to enclose the RNAP cleft. The βGL, $\sigma_{1.2}$ and $\sigma_{2.1}$ (residues 373-396) contact the ntDNA strand of the discriminator from positions -8 to -6 (**Fig. 2C, SMovie 3**); consistently, the βGL deletion destabilizes the RPo and shifts the TSS to the -3A position [22].

The RNAP and *rrnB*P1 complex starts RNA synthesis at the position 9 bp downstream from the -10 element (+1A), which requires DNA scrunching in the open complex [23].

Nucleotide substitutions in several promoter positions, including the discriminator region, were shown to shift the TSS at *rrnB*P1 at position 6 bp downstream from the -10 element (-3A) and stabilize the complex, making it less sensitive to DksA/ppGpp. The RPo structure positioned +1A tDNA at the active site and revealed the mechanism of DNA scrunching, in which the G-7 base of tDNA fits into a pocket surrounded by the β'lid, σfinger (σ region 3.2) and C-6 base (**Fig. 2D, SMovie 3**). The importance of the G-7 base for the DNA scrunching is underscored by its conservation in all seven rRNA promoters in *E. coli* and some rRNA promoters in other proteobacteria (**SFig. 1**). Highly conserved D256 (β'lid) and F522 (σfinger) residues form a salt

bridge and Van der Waals interaction with the G-7 base, respectively. Consistently, we found that an alanine substitution of residue D256 (β'lid) significantly stabilizes RNAP complexes with the *rrnB*P1 promoter (promoter complex half-life $t_{1/2}$ of 135±16 s *vs.* 34±7 s for wild-type RNAP) (**Figs. 3A and B**). The σfinger deletion [24] or the G-7C substitution [23] was shown to shift the TSS to the -3A position, likely eliminating the open complex scrunching.

Open complex scrunching also facilitates the promoter escape of RNAP by reducing abortive RNA synthesis [23]. Robust RNAP escape is required to initiate rRNA transcription approximately every second from each of the seven operons to meet the demand for ribosome synthesis in rapidly dividing *E. coli* [3]. Compared with RPo-*rpsT*P2 containing nonscrunched tDNA [13], RPo-*rrnB*P1 shifts the σfinger ~5 Å away from tDNA, allowing accommodation of one additional base of RNA before its 5'-end reaches the σfinger (**Fig. 2E**). Since the σfinger is the major obstacle to promoter escape [25-27], the partially displaced σfinger in the rRNA promoter RPo due to open complex scrunching may decrease abortive RNA synthesis, promoting the robust expression of rRNA.

Cryo-EM structures of the RNAP and *rrnB*P1 promoter complex with DksA/ppGpp (RP-DksA/ppGpp)

To reveal how DksA/ppGpp binding to RNAP downregulates rRNA transcription, we visualized the RNAP, *rrnB*P1 and DksA/ppGpp complex (RP-DksA/ppGpp) by cryo-EM (**STable 1, SFig. 6**). The classification of the cryo-EM data gave rise to two structures that differed mainly within the RNAP cleft; the first class shows the globular density corresponding to $\sigma_{1.1}$ (class I, RP1-DksA/ppGpp), and the second class shows the right-handed helical density corresponding to the downstream DNA (class II, RP2-DksA/ppGpp) (**Figs. 4A and B, SMovie 4**). In addition, the positions of βlobe/Si1 are different in these classes (**Fig. 4C**).

Both classes show ppGpp binding at sites 1 and 2 and DksA binding at the RNAP secondary channel, as observed in a previous X-ray crystallography study [28]. DksA binds RNAP with its globular domain (G domain, contacts with the β'rim helix), coiled-coil tip (CC tip, contacts with the active site), CC (contacts with the bridge helix, the trigger loop and linkers connecting to the β'Si3), and C-terminal α helix (CT-helix, contacts with the β lobe/SI1 domain) (**Fig. 4A, SMovie 4**). The CC of DksA prevents trigger helix formation and blocks NTP entry from the secondary channel, indicating that DksA must be displaced before RNAP can initiate RNA synthesis [28,29].

Both classes show the duplex DNA density from positions -42 to -14 (from the downstream edge of the UP element to the upstream edge of the -10 element) and also show the ssDNA density of the nontemplate strand of the -10 element (**SFig. 4**). RP1-DksA/ppGpp retains $\sigma_{1.1}$ in the RNAP cleft, indicating that it represents an early stage intermediate during the RPc to RPo transition. While the transcription bubble is likely partially open in RP1-DksA/ppGpp, the density of ntDNA from -5 to +20 and of tDNA from -13 to +20 is not traceable. Analysis of RP1-DksA/ppGpp reveals a DksA/ppGpp-induced conformational change in βlobe/Si1, β'jaw/Si3 and β'clamp, opening the downstream DNA cleft in RNAP and likely reducing the stability of RPo (**Fig. 4D**). Furthermore, the conformational change in βlobe/Si1 establishes a new contact with the DksA CT-helix (**Fig. 4C, SMovie 4**); the deletion of βSi1 reduces the DksA affinity to RNAP and impairs its function [30]. Alanine substitution of an aspartate residue in the CT-helix directly involved in this interaction (D137A) decreases *rrnB*P1 inhibition by DksA both in the absence and in the presence of ppGpp (**Fig. 3C**).

The RP2-DksA/ppGpp complex contains downstream DNA (from +3 to +20) within the RNAP cleft, but the density of the DNA bubble (from -8 to +2) is not traceable (**Fig. 4B, SFig. 4**), suggesting that it represents a late stage intermediate before forming the RPo. The $\sigma_{1.1}$ density is not traceable due to its ejection from the RNAP cleft. The conformations of βlobe/Si1 and β'clamp are akin to the RPo conformation, and the CT-helix of DksA does not contact with the βlobe/Si1 (**Fig. 4C**). Therefore, the transition between the two complexes may reduce the DksA affinity to RNAP and trigger its dissociation, which is an obligatory process to initiate RNA synthesis [28,29].

Conformational change in the σ domain 2 is coupled to $\sigma_{1.1}$ ejection during RPc formation

The RPc structure revealed a significant conformational change in the σ domain 2 (from σ regions 1.2-2.4 including σ nonconserved region (σ_{NCR}, residues 128-372)) comparison with the apo-form holoenzyme RNAP [31] or the RPo containing *rrnB*P1 (this study). Particularly, $\sigma_{1.2}/\sigma_{NCR}$ undergo a rigid rotation toward the clamp to establish contact with the β'clamp-toe (β'CT, residues 143 to 180) (**RPc, Fig. 5A**). Although this interaction was not observed in any previous structural study, it was predicted based on the biochemical/genetic analysis of RNAP promoter escape and early elongation pausing [32]. It was shown that the interaction of the σ_{NCR} and β'CT is important for

196 promoter escape and hinders early elongation pausing, and amino acid substitutions at the interface
197 modulate both processes (**SFig. 7**).
198 The βGL contacts $\sigma_{1.1}$ and $\sigma_{1.2}$ to enclose the RNAP cleft in the apo-form RNAP, which prevents
199 DNA loading (**apo, Fig. 5A**), but the same interaction in the RPo stabilizes the open complex
200 bubble (**Fig. 2B**). In the case of RPc, the $\sigma_{1.2}/\sigma_{NCR}$ rotation disrupts the βGL and σ contact and
201 widens the gap that allows discriminator DNA to enter the RNAP cleft for open complex bubble
202 formation (**RPc, Fig. 5A**). Compared with the apo-form RNAP, the σ_{NCR} and β'CT interaction in
203 the RPc closes the β'clamp, resulting in the ejection of $\sigma_{1.1}$ from the RNAP cleft due to the steric
204 clash between the β'clamp and $\sigma_{1.1}$ (**Fig. 5B**).
205 σ_{NCR} contains a highly negatively charged region (acidic loop, residues 167-213) (**SFig. 8A**). Its
206 conformation has not been determined due to its dynamic behavior, but since it is located near $\sigma_{2.3}$,
207 it seems to prevent nonspecific DNA binding to $\sigma_{2.3}$ (**SFig. 8B**). We speculate that after RNAP
208 recognizes the UP and -35 elements, loading of the -10 element DNA onto σ domain 2 triggers
209 σ_{NCR} rotation due to charge-charge repulsion. After DNA unwinds around the -10 element, σ_{NCR}
210 returns to its position, as seen in the RPo akin to the apo-form RNAP, and may enhance the
211 electrostatic interaction between σ_2 and -10 element DNA (**SFig. 8C**). Consistently, deletion of
212 the acidic loop ($D\sigma_{AL}$) had a weak destabilizing effect on the *rrnB*P1-RNAP complex, without
213 strong effects on DksA inhibition (**Fig. 3**).
214 DksA/ppGpp binding to RNAP also partially opens the DNA loading gate by moving the βlobe/Si1
215 away from $\sigma_{1.1}/\sigma_{1.2}$, but it is not coupled to the $\sigma_{1.1}$ ejection from the RNAP cleft (**R_DksA, Fig.**
216 **5A**). Similarly, the structures of the RNAP-TraR complex and several RNAP-DNA complex
217 intermediates prepared in the presence of TraR also showed the opening of the DNA loading gate
218 by shifting the βlobe/Si1 position but did not show $\sigma_{1.2}/\sigma_{NCR}$ rotation or $\sigma_{1.1}$ ejection from the
219 RNAP cleft (**SFig. 9A**) [13,31].
220 To understand the role of $\sigma_{1.1}$ in rRNA transcription, we characterized an RNAP derivative lacking
221 $\sigma_{1.1}$ ($\Delta\sigma_{1.1}$-RNAP) in terms of its *rrnB*P1 transcription activity and sensitivity to DksA. Compared
222 to the wild-type (WT) RNAP, $\Delta\sigma_{1.1}$-RNAP increases *rrnB*P1 complex stability, both in the absence
223 of DksA (increase $t_{1/2}$ from 34 s to 115 s) and in its presence (increase $t_{1/2}$ from <<10 s to 20 s)
224 (**Figs. 3A and B**), and it decreases sensitivity to DksA (**Fig. 3C**). The results indicate that $\sigma_{1.1}$
225 plays an important role in rRNA transcription and its regulation by DksA/ppGpp.

226

227 **Mechanism of rRNA-specific transcription inhibition by DksA/ppGpp**

228 Structural and biochemical studies of bacterial RNAP transcription suggest that the order of DNA
229 loading around the TSS and DNA opening during promoter recognition may be interchangeable
230 (i.e., DNA melts first outside RNAP (melt-load) or DNA melts after loading inside the RNAP cleft
231 (load-melt)) depending on σ factors, promoters, transcription factors and conditions [26,33]. By
232 combining structural and biochemical data from this and previous studies, we propose two
233 pathways of RPo formation (**Fig. 6, SMovie 5**). We hypothesize that RNAP uses alternative
234 mechanisms of RPo formation, requiring opening of the DNA loading gate (disrupting the βGL
235 contact to σ), $\sigma_{1.1}$ ejection from the DNA binding channel, and unwinding of the -10 element plus
236 discriminator DNA, depending on the absence or presence of DksA/ppGpp.
237 Without DksA/ppGpp (**top, RPo formation**), free RNAP (**R**) binds promoter DNA (**RPc**), which
238 opens the DNA loading gate by ejecting $\sigma_{1.1}$ from the RNAP cleft, making RNAP competent for
239 melting and loading discriminator DNA (**RPi**) into the RNAP cleft, which results in efficient RPo
240 formation. The scrunched open complex (**RPo**) releases RNAP from the rRNA promoter rapidly
241 to proceed with RNA synthesis (**EC**).
242 In the presence of DksA/ppGpp (**bottom, RPo formation with DksA/ppGpp**), DksA/ppGpp
243 binding to RNAP rotates the βlobe/Si1 to DksA, which partially opens the DNA loading gate by
244 disrupting the interaction between GL and σ (**R-DksA**). However, $\sigma_{1.1}$ ejection is uncoupled from
245 RPc formation (**RPc-DksA**), and $\sigma_{1.1}$ remains inside the RNAP cleft until late stages of the open
246 complex formation (**RPi-DksA**). This pathway favors the melt-load model for RPo formation
247 (**RPo-DksA**), in which DNA is accommodated above the βlobe domain and unwinds outside the
248 RNAP cleft (**SFig. 9B**) followed by single-stranded tDNA entry into the active site of RNAP [13].
249 DNA unwinding outside the RNAP cleft is unfavorable in the DksA/ppGpp-free RNAP due to a
250 steric clash of the discriminator DNA with the βlobe. The progression of DNA unwinding from
251 the -10 element to the TSS is energetically less favorable for DksA/ppGpp-sensitive promoters
252 (e.g., *rrnB*P1 and *rpsT*P2) containing the G+C rich discriminator than for less DksA/ppGpp-
253 sensitive promoters (e.g., T7A1 and RNA1) containing an A+T rich discriminator (**SFig. 10**). *E.*
254 *coli* promoters that are sensitive to DksA/ppGpp contain G+C rich discriminators [34]. Replacing the
255 A+T rich discriminator of the *uspA* promoter, which is positively regulated by DksA/ppGpp, with
256 the one from the *rrnB*P1 promoter makes the *uspA* hybrid promoter sensitive to DksA/ppGpp [35],

indicating that discriminator sequences play an important role in responding to DksA/ppGpp. Although DksA could inhibit transcription regardless of the promoter bound to RNAP, by inhibiting NTP entry and folding trigger helix, stable promoter complex formation decreases DksA binding to RNAP, thus relieving the inhibition [28,29]. The completion of discriminator DNA loading into the RNAP cleft can likely occur not only in DksA/ppGpp-insensitive promoters but also in a fraction of the rRNA promoter complexes to maintain a basal level of rRNA expression under stress growth conditions. This likely pushes the βlobe/Si1 away from the CT-helix of DksA (**RPo-DksA**), allowing rapid dissociation of DksA from the RNAP secondary channel (**RPo**) followed by initiate transcription (**EC**).

In this study, we revealed two alternative pathways for opening the gate of the DNA binding channel depending on the absence or presence of DksA/ppGpp (**Fig. 6, SMovie 5**) and shed light on the functions of $\sigma_{1.1}$, $\sigma_{1.2}$, σ_{NCR} and βlobe/Si1 domains to explain how DksA/ppGpp specifically inhibits rRNA transcription. Intriguingly, DksA/ppGpp is able to activate transcription at some σ^{70}-promoters [36] and promoters recognized by alternative σ factors, including σ^{S} [37] and σ^{E} [38]. Neither σ^{S} nor σ^{E} contains $\sigma_{1.1}$ or σ_{NCR}, and the σ^{S} and σ^{E} holoenzymes use the βGL to close the DNA loading gate [39,40]. DNA binding to the σ domain 2 of σ^{S} or σ^{E} cannot facilitate opening of the DNA loading gate, as described in the case of the σ^{70} holoenzyme (**Fig. 6 top**). However, DksA/ppGpp binding followed by movement of the βlobe/Si1 domain could still open the DNA loading gate of these RNAP holoenzymes as described above (**Fig. 6 bottom**), possibly explaining the stimulatory effects of DksA/ppGpp on transcription from some σ^{S}- and σ^{E}-dependent promoters. Further structural analyses of the σ^{70}, σ^{S} and σ^{E} RNAP promoter complex with DksA/ppGpp, together with biochemical characterization of RNAP and promoter functions, will be needed to complete our understanding of DksA/ppGpp-dependent transcription regulation.

REFERENCES

Uncategorized References

1 Nomura, M. Regulation of ribosome biosynthesis in Escherichia coli and Saccharomyces cerevisiae: diversity and common principles. *J Bacteriol* **181**, 6857-6864, (1999).

2 Zengel, J. M. & Lindahl, L. Diverse mechanisms for regulating ribosomal protein synthesis in Escherichia coli. *Prog Nucleic Acid Res Mol Biol* **47**, 331-370, (1994).

3 Bremer, H. & Dennis, P. P. Modulation of Chemical Composition and Other Parameters of the Cell at Different Exponential Growth Rates. *EcoSal Plus* **3**, (2008).

4 Gourse, R. L. *et al.* Transcriptional responses to ppGpp and DksA. *Annu Rev Microbiol* **72**, 163-184, (2018).

5 Ross, W. *et al.* A third recognition element in bacterial promoters: DNA binding by the alpha subunit of RNA polymerase. *Science* **262**, 1407-1413, (1993).

6 Winkelman, J. T. & Gourse, R. L. Open complex DNA scrunching: A key to transcription start site selection and promoter escape. *Bioessays* **39**, (2017).

7 Gaal, T., Bartlett, M. S., Ross, W., Turnbough, C. L., Jr. & Gourse, R. L. Transcription regulation by initiating NTP concentration: rRNA synthesis in bacteria. *Science* **278**, 2092-2097, (1997).

8 Cashel, M. & Gallant, J. Two compounds implicated in the function of the RC gene of Escherichia coli. *Nature* **221**, 838-841, (1969).

9 Potrykus, K. & Cashel, M. (p)ppGpp: Still Magical? *Annual Review of Microbiology* **62**, 35-51, (2008).

10 Hauryliuk, V., Atkinson, G. C., Murakami, K. S., Tenson, T. & Gerdes, K. Recent functional insights into the role of (p)ppGpp in bacterial physiology. *Nat Rev Microbiol* **13**, 298-309, (2015).

11 Murakami, K. S., Masuda, S., Campbell, E. A., Muzzin, O. & Darst, S. A. Structural basis of transcription initiation: an RNA polymerase holoenzyme-DNA complex. *Science* **296**, 1285-1290, (2002).

12 Zhang, Y. *et al.* Structural basis of transcription initiation. *Science* **338**, 1076-1080, (2012).

13 Chen, J. *et al.* Stepwise Promoter Melting by Bacterial RNA Polymerase. *Mol Cell*, (2020).

14 Newlands, J. T., Ross, W., Gosink, K. K. & Gourse, R. L. Factor-independent activation of Escherichia coli rRNA transcription. II. characterization of complexes of rrnB P1 promoters containing or lacking the upstream activator region with Escherichia coli RNA polymerase. *J Mol Biol* **220**, 569-583, (1991).

15 Murakami, K., Kimura, M., Owens, J. T., Meares, C. F. & Ishihama, A. The two alpha subunits of Escherichia coli RNA polymerase are asymmetrically arranged and contact different halves of the DNA upstream element. *Proc Natl Acad Sci U S A* **94**, 1709-1714, (1997).

16 Fujita, N., Endo, S. & Ishihama, A. Structural requirements for the interdomain linker of alpha subunit of Escherichia coli RNA polymerase. *Biochemistry* **39**, 6243-6249, (2000).

17 Sreenivasan, R. *et al.* Fluorescence Resonance Energy Transfer Characterization of DNA Wrapping in Closed and Open Escherichia coli RNA Polymerase-lambdaP(R) Promoter Complexes. *Biochemistry* **55**, 2174-2186, (2016).

18 Doniselli, N. *et al.* New insights into the regulatory mechanisms of ppGpp and DksA on Escherichia coli RNA polymerase-promoter complex. *Nucleic acids research* **43**, 5249-5262, (2015).

19 Feklistov, A. & Darst, S. A. Structural basis for promoter-10 element recognition by the bacterial RNA polymerase sigma subunit. *Cell* **147**, 1257-1269, (2011).

20 Murphy, H. & Cashel, M. Isolation of RNA polymerase suppressors of a (p)ppGpp deficiency. *Methods Enzymol* **371**, 596-601, (2003).

21 Rutherford, S. T., Villers, C. L., Lee, J. H., Ross, W. & Gourse, R. L. Allosteric control of Escherichia coli rRNA promoter complexes by DksA. *Genes Dev* **23**, 236-248, (2009).

22 NandyMazumdar, M. *et al.* RNA polymerase gate loop guides the nontemplate DNA strand in transcription complexes. *Proc Natl Acad Sci U S A* **113**, 14994-14999, (2016).

23 Winkelman, J. T., Chandrangsu, P., Ross, W. & Gourse, R. L. Open complex scrunching before nucleotide addition accounts for the unusual transcription start site of E. coli ribosomal RNA promoters. *Proc Natl Acad Sci U S A* **113**, E1787-1795, (2016).

24 Pupov, D., Petushkov, I., Esyunina, D., Murakami, K. S. & Kulbachinskiy, A. Region 3.2 of the sigma factor controls the stability of rRNA promoter complexes and potentiates their repression by DksA. *Nucleic acids research* **46**, 11477-11487, (2018).

25 Murakami, K. S., Masuda, S. & Darst, S. A. Structural basis of transcription initiation: RNA polymerase holoenzyme at 4 A resolution. *Science* **296**, 1280-1284, (2002).

26 Mazumder, A. & Kapanidis, A. N. Recent Advances in Understanding sigma70-Dependent Transcription Initiation Mechanisms. *J Mol Biol* **431**, 3947-3959, (2019).

27 Kulbachinskiy, A. & Mustaev, A. Region 3.2 of the sigma subunit contributes to the binding of the 3'-initiating nucleotide in the RNA polymerase active center and facilitates promoter clearance during initiation. *J Biol Chem* **281**, 18273-18276, (2006).

28 Molodtsov, V. *et al.* Allosteric Effector ppGpp Potentiates the Inhibition of Transcript Initiation by DksA. *Molecular cell* **69**, 828-839 e825, (2018).

29 Stumper, S. K. *et al.* Delayed inhibition mechanism for secondary channel factor regulation of ribosomal RNA transcription. *Elife* **8**, (2019).

30 Parshin, A. *et al.* DksA regulates RNA polymerase in Escherichia coli through a network of interactions in the secondary channel that includes Sequence Insertion 1. *Proc Natl Acad Sci U S A* **112**, E6862-6871, (2015).

31 Chen, J. *et al.* E. coli TraR allosterically regulates transcription initiation by altering RNA polymerase conformation. *Elife* **8**, (2019).

32 Leibman, M. & Hochschild, A. A sigma-core interaction of the RNA polymerase holoenzyme that enhances promoter escape. *EMBO J* **26**, 1579-1590, (2007).

33 Glyde, R. *et al.* Structures of Bacterial RNA Polymerase Complexes Reveal the Mechanism of DNA Loading and Transcription Initiation. *Molecular cell* **70**, 1111-1120 e1113, (2018).

34 Sanchez-Vazquez, P., Dewey, C. N., Kitten, N., Ross, W. & Gourse, R. L. Genome-wide effects on Escherichia coli transcription from ppGpp binding to its two sites on RNA polymerase. *Proc Natl Acad Sci U S A* **116**, 8310-8319, (2019).

35 Gummesson, B., Lovmar, M. & Nystrom, T. A proximal promoter element required for positive transcriptional control by guanosine tetraphosphate and DksA protein during the stringent response. *J Biol Chem* **288**, 21055-21064, (2013).

36 Paul, B. J., Berkmen, M. B. & Gourse, R. L. DksA potentiates direct activation of amino acid promoters by ppGpp. *Proc Natl Acad Sci U S A* **102**, 7823-7828, (2005).

37 Girard, M. E. *et al.* DksA and ppGpp Regulate the sigma(S) Stress Response by Activating Promoters for the Small RNA DsrA and the Anti-Adapter Protein IraP. *Journal of bacteriology* **200**, (2018).

38 Costanzo, A. *et al.* ppGpp and DksA likely regulate the activity of the extracytoplasmic stress factor sigmaE in Escherichia coli by both direct and indirect mechanisms. *Mol Microbiol* **67**, 619-632, (2008).

39 Li, L., Fang, C., Zhuang, N., Wang, T. & Zhang, Y. Structural basis for transcription initiation by bacterial ECF sigma factors. *Nat Commun* **10**, 1153, (2019).

40 Xu, J. *et al.* Crl activates transcription by stabilizing active conformation of the master stress transcription initiation factor. *Elife* **8**, (2019).

41 Murakami, K. S. X-ray crystal structure of *Escherichia coli* RNA polymerase sigma70 holoenzyme. *J Biol Chem* **288**, 9126-9134, (2013).

42 Zivanov, J. *et al.* New tools for automated high-resolution cryo-EM structure determination in RELION-3. *Elife* **7**, (2018).

43 Zheng, S. Q. *et al.* MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat Methods* **14**, 331-332, (2017).

44 Zhang, K. Gctf: Real-time CTF determination and correction. *Journal of structural biology* **193**, 1-12, (2016).

45 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).

46 Pettersen, E. F. *et al.* UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605-1612, (2004).

47 Afonine, P. V. *et al.* Joint X-ray and neutron refinement with phenix.refine. *Acta Crystallogr D Biol Crystallogr* **66**, 1153-1163, (2010).

48 Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126-2132, (2004).

49 Svetlov, V. & Artsimovitch, I. Purification of bacterial RNA polymerase: tools and protocols. *Methods Mol Biol* **1276**, 13-29, (2015).

50 SantaLucia, J., Jr. A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proc Natl Acad Sci U S A* **95**, 1460-1465, (1998).

51 Dienemann, C., Schwalb, B., Schilbach, S. & Cramer, P. Promoter Distortion and Opening in the RNA Polymerase II Cleft. *Mol Cell* **73**, 97-106 e104, (2019).

METHODS

Purification of *E. coli* WT RNAP, RNAP derivatives, and DksA. *E. coli* σ^{70} RNAP holoenzyme and DksA were purified as described previously [28,41]. *E. coli* RNAP derivatives were purified by the same method as WT RNAP.

Preparation of *rrnB*P1 DNA. The *rrnB*P1 promoter DNA was synthesized (IDT) according to the native *rrnB*P1 sequence and annealed in a 40 µL reaction mixture containing 10 mM Tris-HCl (pH 8.0), 50 mM NaCl, and 1 mM EDTA to a final concentration of 0.5 mM. The solution was

heated at 95 °C for 10 min, and then the temperature was gradually decreased to 22 °C. The sequence of the nontemplate strand is 5'-CAGAAAATTATTTTAAATTTCCTCTTGTCAGGCCGGAATAACTCCCTATAATGCGCC ACCACTGACACGGACTCTACGAG-3'. The transcription start site is underlined, and the template sequence is 5'-CTCGTAGAGTCCGTGTCAGTGGTGGCGCATTATAGGGAGTTATTCCGGCCTGACAAG AGGAAATTTAAAATAATTTTCTG -3'.

Cryo-EM sample preparation. To prepare the RNAP and *rrnB*P1 promoter complex, *E. coli* σ^{70} RNAP (20 μM) and *rrnB*P1 promoter DNA (40 μM) were preincubated for 5 min at 37 °C in buffer (10 mM HEPES, pH 8.0, 50 mM NaCl, 0.1 mM EDTA, 5 mM DTT and 5 mM $MgCl_2$) followed by the addition of ATP and a nonhydrolyzable CMPCPP (cytidine-5'-[(α,β)-methyleno]triphosphate, Jena Bioscience) (2 mM each). After mixing, the reaction was further incubated for 5 mins at 37 °C.

To prepare RP-DksA/ppGpp, *E. coli* σ^{70} RNAP (20 μM) was preincubated with a 5-fold molar excess of DksA (100 μM) and ppGpp (2 mM) for 5 min at 37 °C in buffer (10 mM HEPES, pH 8.0, 50 mM NaCl, 0.1 mM EDTA, 5 mM DTT and 5 mM $MgCl_2$). The *rrnB*P1 promoter DNA (40 μM) was added to the reaction and further incubated for 5 mins at 37 °C. Before freezing the grids, 8 mM CHAPSO (Hampton research) was added to the reaction. A 3.5 μL sample was applied to a glow-discharged C-Flat Holey Carbon grid (Cu 2/1, 400 mesh), blotted and plunge-frozen in liquid ethane using a Vitrobot Mark IV (FEI, USA) with 95% humidity at 4 °C.

Cryo-EM data acquisition. The grid was imaged using a 300 keV Titan Krios (Thermo Fisher) microscope equipped with a K3 direct electron detector (Gatan) and controlled by the Latitude S (Gatan, Inc.) software at the National Cancer Institute's Cryo-EM Facility at Frederick. The defocus range was 1.0 – 3.0 μm, and the magnification was 81,000X in electron counting mode (pixel size = 1.08 Å/pixel). Forty frames per movie were collected with a dose of 1.125 $e^-/Å^2$/frame, giving a total dose of 45 $e^-/Å^2$.

Cryo-EM data processing. The RNAP-*rrnB*P1 complex with ATP/CMPCPP data was processed using Relion3.0.8 [42]. A total of 8,315 movies were collected, aligned and dose weighted using

MotionCor2 [43]. CTF fitting was performed with Gctf [44]. Initially, approximately 1,000 particles were manually picked to generate particle templates followed by automated picking, resulting in a total of 1,449,010 particles subjected to 2D classification. From the 2D classes, 1,442,810 particles were chosen for the 3D classification to 4 classes. Poorly populated classes were removed, resulting in datasets of 541,257 (37%) particles for the first class (RPc) and 464,512 (32%) particles for the second class. The first class was further 3D classified without alignments twice to further clean the data, resulting in datasets of 67,187 particles. The particles were refined and postprocessed to generate the density map at 4.14 Å resolution. The resolution of the density map of the second class was 3.53 Å.

The RNAP-*rrnB*P1 complex data were processed using Relion3.0.8. A total of 4748 movies were collected, aligned and dose weighted using MotionCor2. CTF fitting was performed with Gctf. Approximately 1,000 particles were manually picked to generate particle templates followed by automated picking, resulting in a total of 563,500 particles. Particles were 2D classified, and 561,753 particles were chosen for the 3D classification. Of the four 3D classes, class 1 (RPo) was the most populated class (349,752 particles, 62%) and was autorefined. The map was postprocessed to give a structure of RPo at 3.53 Å.

The RP-DksA/ppGpp complex data were processed using cryoSPARC V2.9.0 [45]. A total of 4,926 movies were collected, and the movies were aligned, and dose weighted using Patch-motion correction. CTF fitting was performed with Patch-CTF estimation. Initially, approximately 1,000 particles were manually picked to generate particle templates followed by automated picking, resulting in a total of 418,049 particles subjected to 2D classification. After two rounds of 2D classification to remove junk particles, 361,048 particles were used to generate two *ab initio* models. Junk particles were removed, resulting in a dataset of 275,629 particles chosen for the 3D classification (heterogenous refinement). Poorly populated classes were removed, resulting in a dataset of 49,995 particles to generate the density map at 3.62 Å resolution for the first class (RP1-DksA/ppGpp) and a dataset of 79,275 particles to generate the density map at 3.58 Å resolution for the second class (RP2-DksA/ppGpp). The particles were 3D autorefined without the mask and postprocessed (homogenous refinement).

Structure refinement. To refine the closed and open complex structures, the *E. coli* RNAP holoenzyme crystal structure (PDB: 4YG2) was manually fit into the cryo-EM density map using

Chimera [46] and real-space refined using Phenix [47]. In the real-space refinement, the domains of RNAP were rigid-body refined and then subsequently refined with secondary structure, Ramachandran, rotamer and reference model restraints. To refine the structures of RP1-DksA/ppGpp and RP2-DksA/ppGpp, *E. coli* RNAP and DksA/ppGpp complex crystal structures (PDB: 5VSW) were manually fit into the cryo-EM density map using Chimera. DNA was manually built by using Coot [48]. The structure was refined by the same method as the closed and open complex structures.

Preparation of RNAP and transcription factors for *in vitro* transcription. Mutant variants of RNAP, σ^{70} and DksA were obtained by site-directed mutagenesis. The D256A substitution in the β′ subunit was obtained in pVS10 encoding all RNAP subunits, with the *rpoC* gene containing a C-terminal His_6-tag [49]. The σ^{70} and DksA variants containing an N-terminal His_6-tag were cloned into pET28. To obtain $\sigma\Delta_{1.1}$, the 5'-terminal part of the *rpoD* gene encoding residues 2-94 was deleted. To obtain $\sigma\Delta_{AL}$, codons 168-212 were replaced with three glycine codons. All proteins were expressed in *E. coli* BL21(DE3). The wild-type and mutant core RNAPs were purified using Polymin P precipitation followed by heparin (HiTrap Heparin column), Ni-affinity (HisTrap HP column) and anion exchange (MonoQ column) chromatography steps (all columns from GE Healthcare) as described previously [49]. The wild-type and mutant σ^{70} factors were purified from inclusion bodies with subsequent renaturation and Ni-affinity chromatography as previously described [27]. The $\sigma\Delta_{1.1}$ protein was subjected to thrombin protease (GE Healthcare) treatment in PBS buffer (10 hours of incubation at 4 °C with 10 units of protease per mg of protein), followed by incubation with Ni-NTA agarose (GE Healthcare) to remove the His-tag and His-tagged thrombin. To purify DksA, a bacterial pellet from 0.5 liters of cell culture was resuspended in 25 ml of lysis buffer (50 mM Tris-HCl, pH 7.9, 250 mM NaCl, 10 mM EDTA, 0.5 mM phenylmethylsulfonyl fluoride, 1 mM 2-mercaptoethanol, 0.1 mM $ZnCl_2$) and lysed using a French press. The supernatant obtained after centrifugation was loaded onto a 5-ml HiTrap chelating column (GE Healthcare) charged with Ni^{2+} and equilibrated with loading buffer (10 mM Tris-HCl, pH 7.9, 500 mM NaCl, 0.5 mM 2-mercaptoethanol, 0.1 mM $ZnCl_2$). The column was washed with the same buffer containing 60 mM imidazole, and DksA was eluted with buffer containing 300 mM imidazole and dialyzed overnight against 50 mM Tris-HCl, 300 mM NaCl, 1 mM DTT, and 0.1 mM $ZnCl_2$. Glycerol was added to 50%, and aliquots were stored at -70 °C.

Transcription *in vitro*. Analysis of transcription *in vitro* was performed using a supercoiled pTZ19 template containing *rrnB*P1 cloned 88 nt upstream of a *his* terminator (Pupov et al., 2018); the second transcript monitored in the assays was 110 nt RNA-I encoded by the *ori* region of the plasmid. For measurements of promoter complex stabilities, promoter complexes were prepared by mixing core RNAP (100 nM final concentration) with wild-type or mutant σ^{70} factors (250 nM) in transcription buffer (40 mM Tris-HCl, pH 7.9, 10 mM $MgCl_2$, 40 mM KCl) and supercoiled plasmid DNA (10 nM), followed by incubation for 7 min at 37 °C. DksA and ppGpp were added at 2 μM and 200 μM, respectively, when indicated. An upstream fork-junction competitor DNA was added (template strand 5'-ACGAGCCGGAAGCAT, nontemplate strand 5'-ATGCTTCCGGCTCGTATAATGTGTGGAA; the -10 sequence is underlined) to 2 μM, and the samples were incubated at 37 °C for the indicated time intervals. NTP substrates were added to final concentrations of 200 μM ATP, CTP, GTP, and 10 μM UTP, with the addition of α-[^{32}P]-UTP together with rifapentin (5 μg/ml) to prevent reinitiation. The reactions were stopped after 5 min with 8 M urea and 20 mM EDTA, and RNA products were separated by 15% denaturing PAGE, followed by phosphor imaging. To calculate the observed half-life times for promoter complex dissociation ($t_{1/2}$), the data were fitted to the one-exponential equation $A = A_0 \times \exp(-t \times k_{obs})$, where A is the RNAP activity at a given time point after competitor addition, A_0 is the activity measured in the absence of the competitor, k_{obs} is the observed rate constant, and $t_{1/2} = \ln 2/k_{obs}$.

For measurements of apparent DksA affinities, promoter complexes were prepared in the same way with 50 nM core RNAP, 250 nM σ^{70} (250 nM) and 2 nM supercoiled plasmid DNA in transcription buffer containing 100 μg/ml BSA for 7 min at 37 °C, followed by the addition of DksA (from 10 nM to 10 μM), either in the absence or in the presence of ppGpp (100 μM). Transcription was performed for 15 min at 37 °C with 200 μM ATP, CTP, GTP, and 10 μM UTP (plus α-[^{32}P]-UTP), and RNA products were analyzed as described above. The apparent dissociation constant values ($K_{d,app}$) were calculated from the hyperbolic equation: $A = A_{max} \times (1 - [DksA]/(K_{d,app} + [DksA]))$, where A is the observed RNAP activity and A_{max} is the RNAP activity measured in the absence of DksA.

DNA duplex free energy calculation. DNA duplex free energies were analyzed based on nearest-neighbor thermodynamics [50,51]. Briefly, the Python script was written to read a sequence from a text file, calculate the DNA duplex free energy of dinucleotides, sum these values over an 8-base window and report these sums for the first base of the central nucleotide of the window (e.g., sum for the first window base 1-8 will be reported for base 4).

Data Availability. The cryo-EM density maps have been deposited in EMDataBank under accession codes EMDB: EMD-21879 (RPc), EMD-21880 (RPo), EMD-21881 (RP1-DksA/ppGpp), and EMD-21883 (RP2-DksA/ppGpp). Atomic coordinates for the reported cryo-EM structures have been deposited with the Protein Data Bank under accession numbers 6WR6, 6WR8, 6WRD and 6WRG.

ACKNOWLEDGMENTS

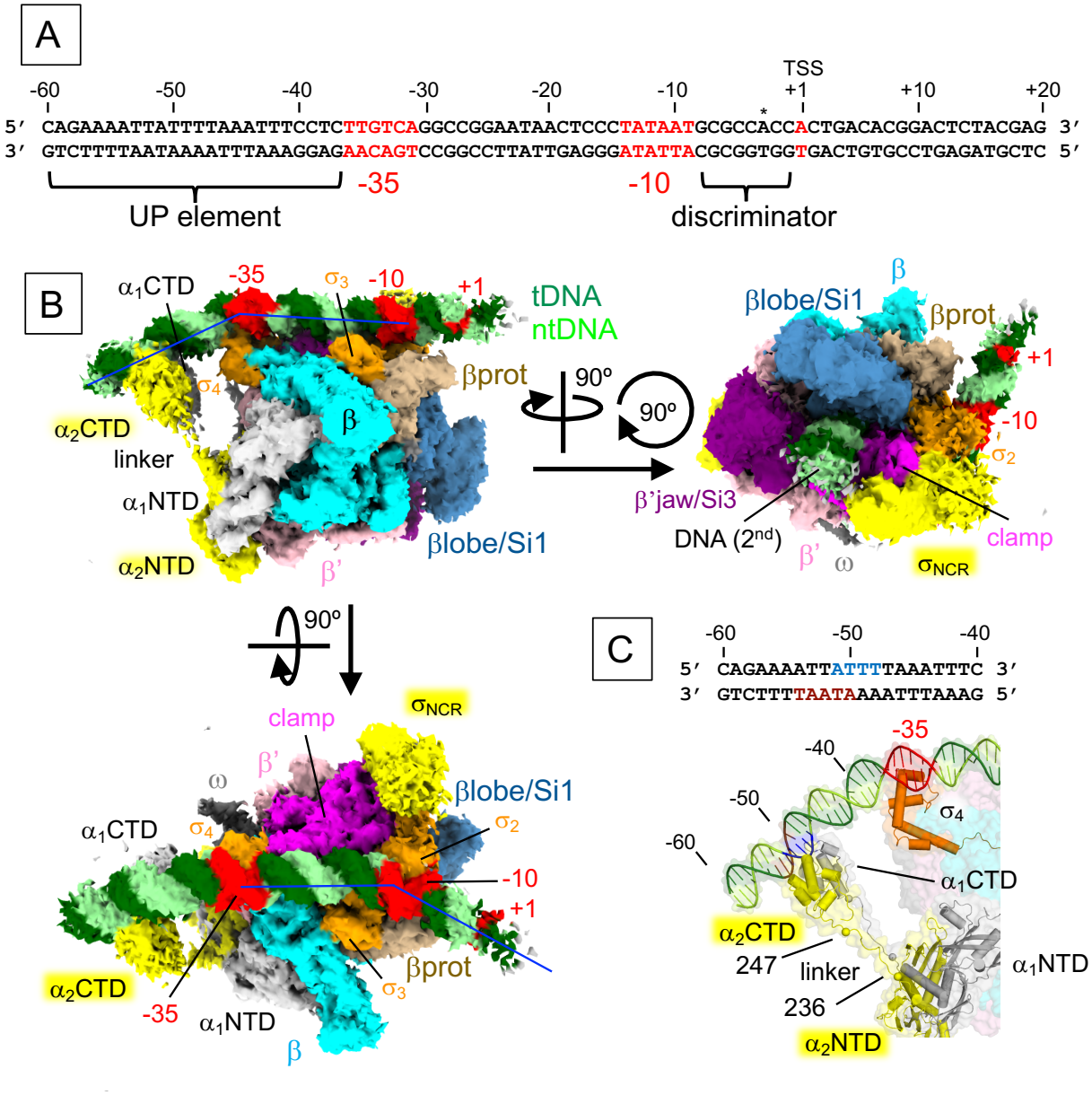
We thank Carol Bator at the Penn State Huck Institute Cryo-EM facility for supporting the cryo-EM data collections. We thank Christian Dienemann at the Max Planck Institute for providing the DNA duplex free energy calculation methodology. This research was, in part, supported by the National Cancer Institute's National Cryo-EM Facility at the Frederick National Laboratory for Cancer Research under contract HSSN261200800001E. This work was supported by NIH grants (R01 GM087350 and R35 GM131860 to K.S.M.) and Russian Science Foundation (19-14-00359) and Russian Foundation for Basic Research (18-34-20095) to A.K.

AUTHOR CONTRIBUTIONS

Y.S. prepared samples and cryo-EM grids. Y.S. and M.Z.Q. collected and processed cryo-EM data. K.S.M. built, refined and validated the structures. D.P. and D.E. performed biochemical assays under supervision of A.K. K.S.M. designed and supervised research. All authors participated in the interpretation of the results and in writing the manuscript.

Competing interest The authors declare no competing interests.

Correspondence and requests for materials should be addressed to K.S.M. or A.K.



564

565 **Figure 1. Cryo-EM structure of the RNAP - *rrnB*P1 closed complex (RPc). A)** The sequence
566 of the *E. coli rrnB*P1 promoter DNA used for cryo-EM. The UP element, -35 element, -10 element,
567 transcription start site (TSS, +1) and discriminator sequence are indicated. An altered TSS from
568 the nonscrunched open complex is indicated by an asterisk. **B)** Orthogonal views of the RPc cryo-
569 EM density map. Subunits and domains of RNAP and DNA are colored and labeled (βprot,
570 βprotrusion; tDNA, template DNA; ntDNA, nontemplate DNA). The density of downstream DNA
571 beyond the +4 position is not traceable. Blue lines denote the direction of the DNA axis, with kinks
572 at approximately -37 and -13. The second DNA at the RNAP cleft is indicated (DNA (2^{nd})). **C)** A
573 magnified view showing the αCTDs and UP element interaction. The domains of α subunits, σ_4
574 and DNA are depicted as ribbon models with a partially transparent surface. At the top, the
575 sequence of the UP element is shown. The ntDNA (-51 to -48) and tDNA (-54 to -50) sequences
576 binding α_1CTD and α_2CTD are highlighted in blue and brown, respectively.

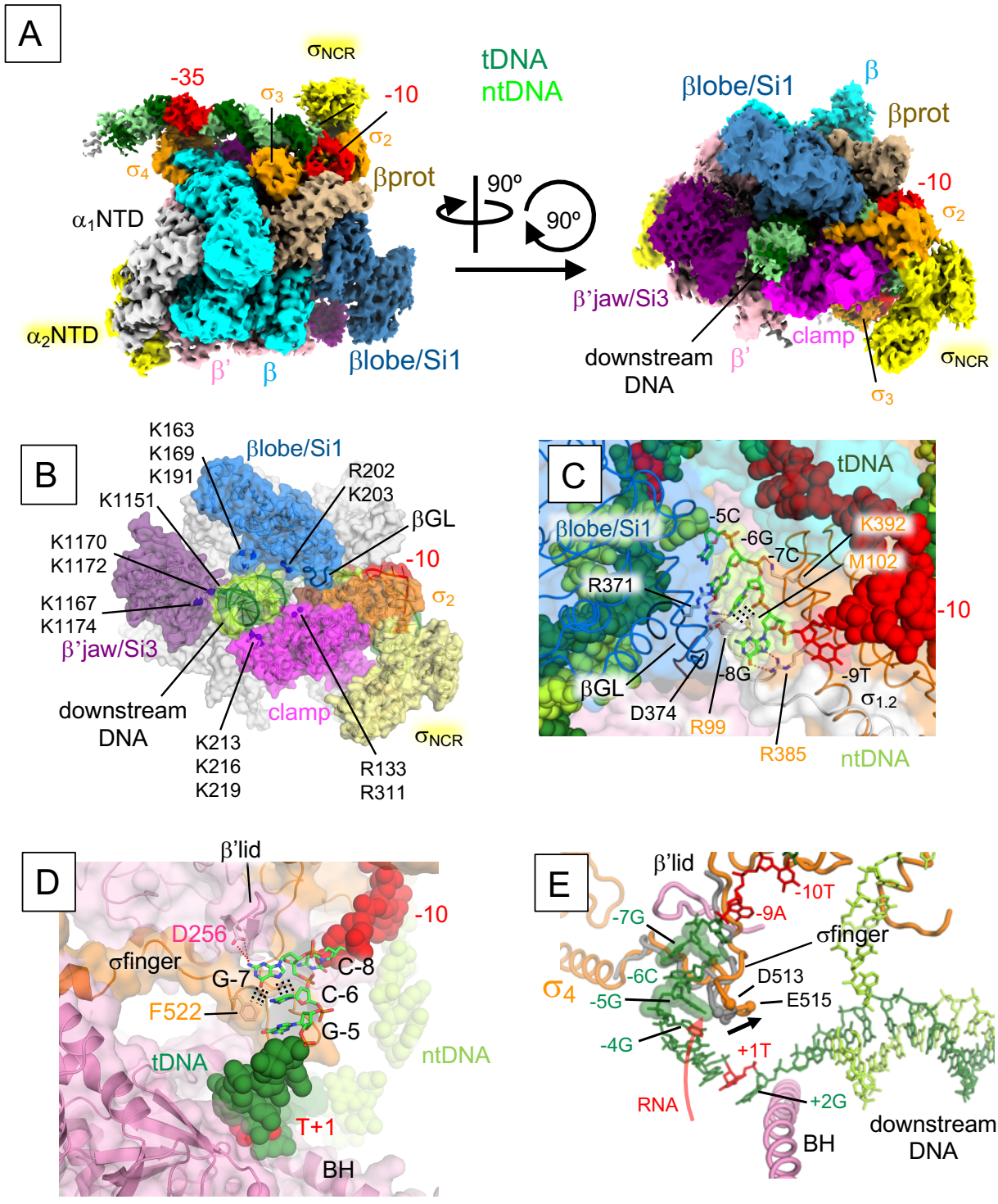
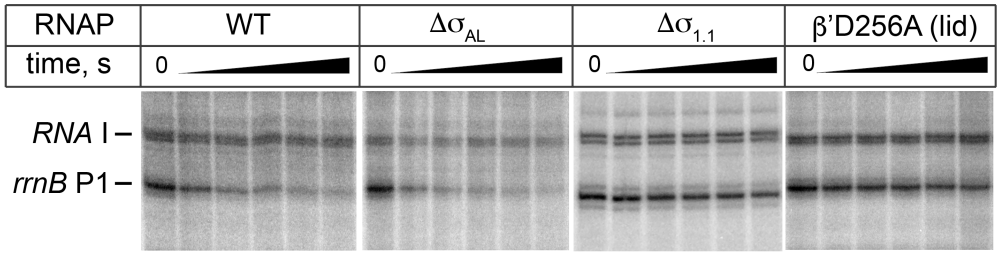


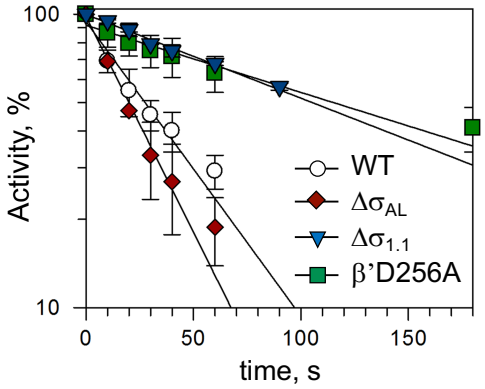
Figure 2. Cryo-EM structure of the RNAP – *rrnB*P1 open complex (RPo). A) Orthogonal views of the RPo cryo-EM density map. Subunits and domains of RNAP and DNA are colored and labeled the same as in Fig. 1B. **B)** The structure of the RPo, highlighting basic residues of the βlobe/Si1 (blue), β'jaw/Si3 (purple) and β'clamp (pink) interacting with downstream DNA (green) to stabilize the RPo. The structure is shown as a ribbon model with a transparent surface, and the basic residues are shown as spheres and labeled. **C)** Close-up view of RNAP (βGL, $\sigma_{1.1}$ and $\sigma_{1.2}$)

584 and discriminator DNA (ntDNA) interaction. β and σ are depicted as ribbon models with
585 transparent surfaces, and DNA is shown as CPK spheres. The G-8, C-7 and G-6 bases (stick model
586 with transparent CPK spheres) that form salt bridges and Van del Waals interactions with residues
587 from the βGL and $\sigma_{1.2}$ (side chains shown as sticks; βGL R371 and D374; $\sigma_{1.2}$ R99 and M102) are
588 shown (depicted by red and black dashed lines). **D)** Close-up view of the RNAP (β'lid and σfinger)
589 and discriminator DNA (tDNA) interaction. The G-7 base inserts into the pocket formed by the
590 β'lid, σfinger and C-6 base. β' and σ are depicted as ribbon models with transparent surfaces, and
591 DNA is shown as a stick model and CPK representation. The residues forming salt bridges and
592 Van del Waals interactions with the G-7 base are shown (depicted by red and black dashed lines).
593 **E)** Comparison of the σfinger in RPo-*rrnB*P1 (this study, orange) and RPo-*rpsT*P2 [13] (gray). The
594 RPo-*rrnB*P1 structure is depicted as cartoon (RNAP) and stick (DNA) models with the σfinger
595 from RPo-*rpsT*P2 (gray). When the *rrnB*P1 tDNA scrunches at the -7G(t) position, -5G(t) is
596 located below the σfinger (orange), which shifts the σfinger position compared to that in
597 nonscrunched RPo. The σfinger dislocation (black arrow, 5 Å at E515) makes additional space for
598 RNA extension (red arrow).
599
600

A



B



RNAP	$t_{1/2}$, s	
	-DksA	**+DksA**
WT	34.4 ± 7.1 **1**	<< 10 s
$\Delta\sigma_{AL}$	20.7 ± 4.6 **0.6**	ND
$\Delta\sigma_{1.1}$	115 ± 19 **3.3**	20 ± 3
β′D256A	135 ± 16 **3.9**	ND

C

RNAP	$K_{d,app}$ (nM)	
	-ppGpp	+ppGpp
WT	2200 ± 430 **1**	860 ± 180 **1**
$\Delta\sigma_{AL}$	3000 ± 510 **1.4**	1150 ± 220 **1.3**
$\Delta\sigma_{1.1}$	>10000 **>5**	ND
β′D256A	4540 ± 150 **2.1**	1640 ± 220 **1.9**
DksA		
D137A	3900 ± 730 **1.8**	≥2500 **≥3**

Figure 3. Stabilities of *rrnB*P1 promoter complexes formed by wild-type and mutant RNAPs and their sensitivities against DksA/ppGpp. A) Sensitivity of *rrnB*P1 promoter complexes to heparin. Preformed promoter complexes were incubated with heparin for the indicated time intervals, followed by the addition of NTPs and rifapentin. **B**) Kinetics of promoter complex dissociation for wild-type and mutant RNAPs. The half-lives of the promoter complexes for each RNAP are shown in the table to the right (mean values and standard deviations from three independent experiments). **C**) Apparent DksA affinities to wild-type and mutant RNAPs on the *rrnB*P1 promoter.

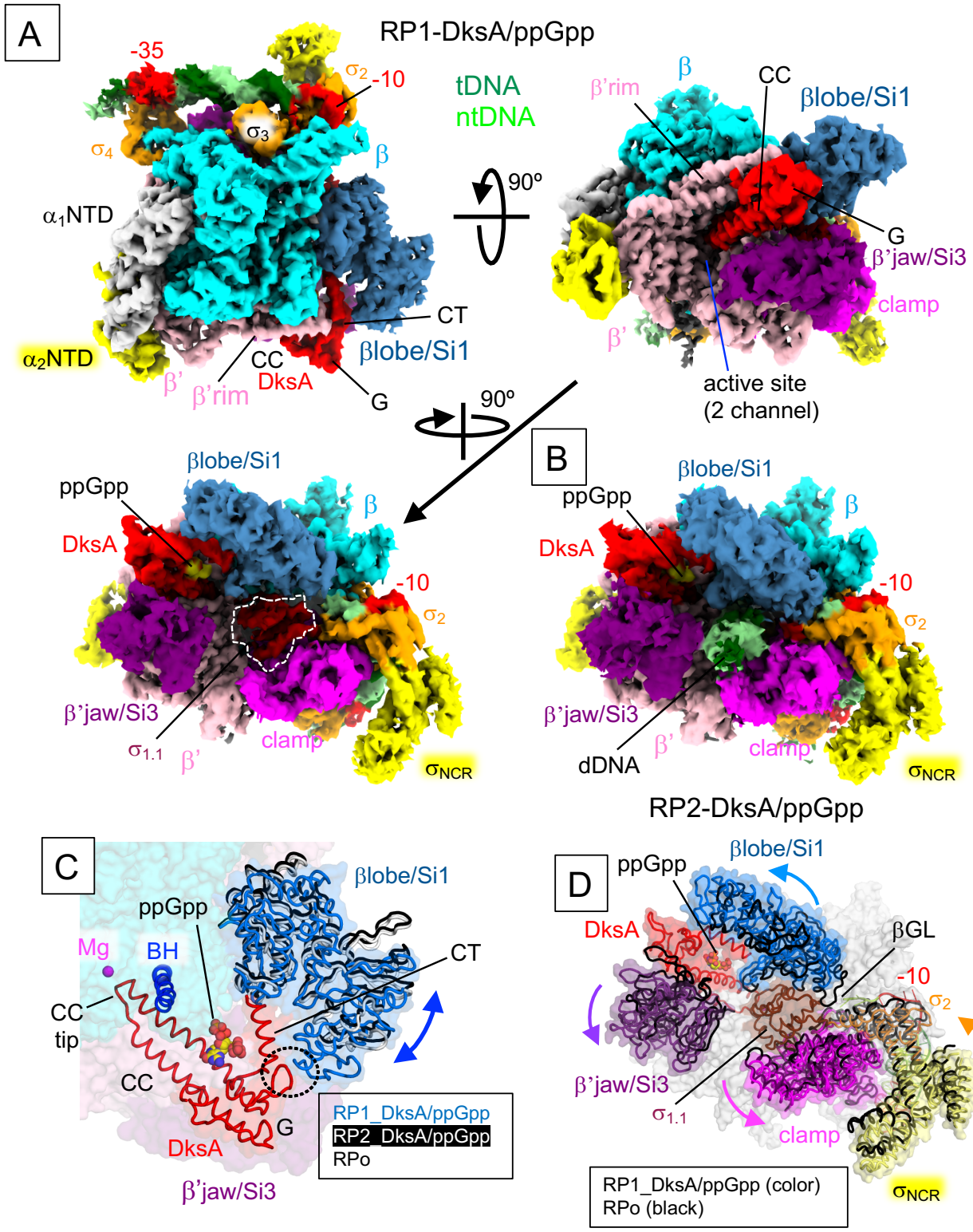


Figure 4. Cryo-EM structures of the RNAP - *rrnB*P1 complex with DksA/ppGpp (RP-DksA/ppGpp). A) Orthogonal views of the RP1-DksA/ppGpp cryo-EM density map. DNA, RNAP and DksA (G, G domain; CC, CC domain; CT, CT-helix) are indicated and colored. $\sigma_{1.1}$ is

highlighted by a white dash. **B)** The RP2-DksA/ppGpp cryo-EM density map. The downstream DNA is accommodated in the RNAP cleft. **C)** Close-up view of the βlobe/Si1 conformational changes upon DksA binding, $\sigma_{1.1}$ ejection and downstream DNA binding. The structures of the βlobe/Si1 in RP1-DksA/ppGpp (light blue), RP2-DksA/ppGpp (white) and RPo (black) are depicted as ribbon models with transparent surfaces and ribbon models (DksA, BH: bridge helix) of RP1-DksA/ppGpp. The interaction between βlobe/Si1 and DksA in RP1-DksA/ppGpp is highlighted by a black dashed oval. **D**) Conformational changes in the RNAP mobile domains upon binding of DksA/ppGpp. The structures of the RNAP mobile domains in RP1-DksA/ppGpp (colored) and RPo (black) are depicted as ribbon models with transparent surfaces of RP1-DksA/ppGpp.

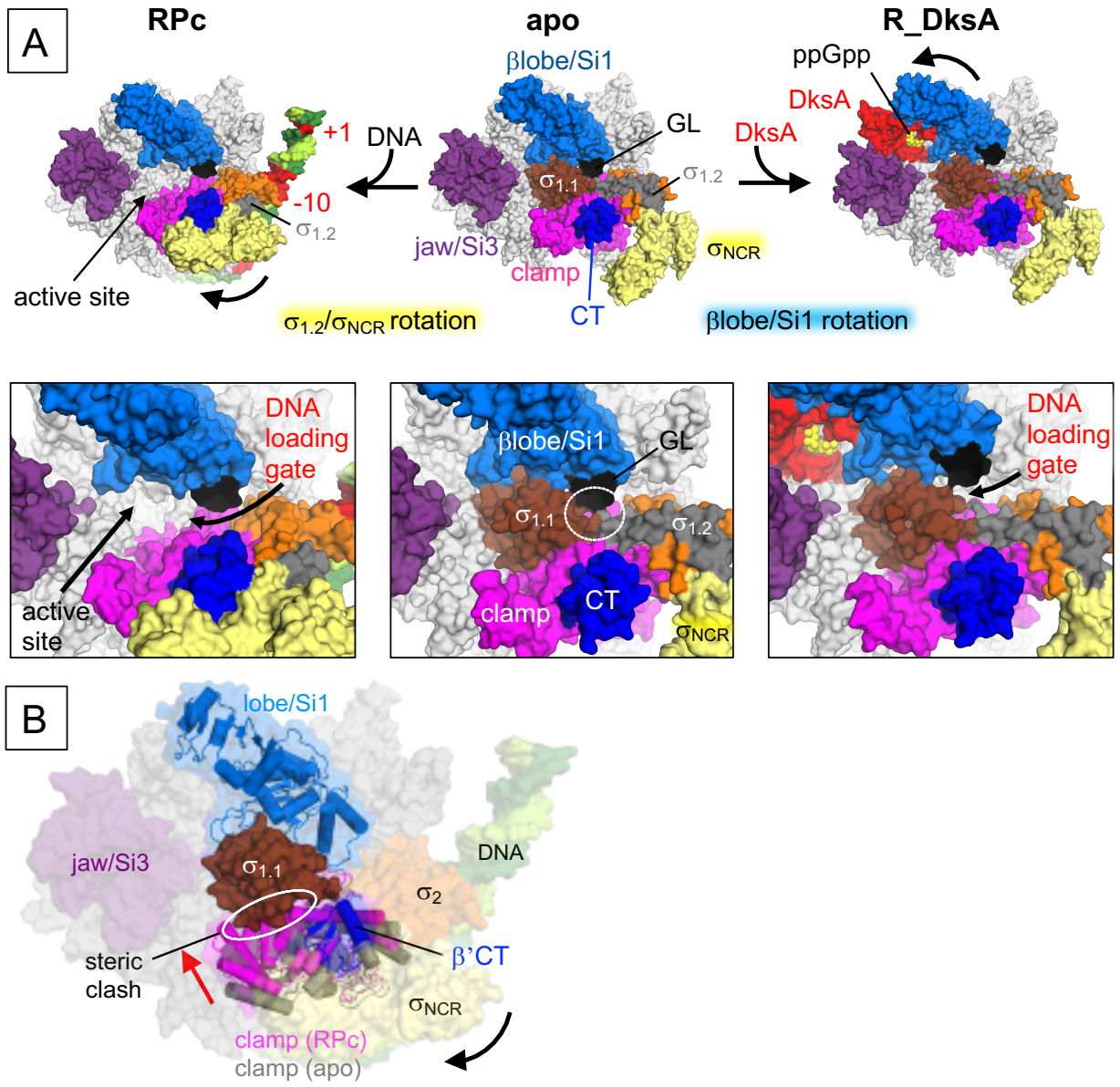


Figure 5. Opening the DNA loading gate by moving $\sigma_{1.2}/\sigma_{NCR}$ or βlobe/Si1 domain. A) Comparison of the $\sigma_{1.2}/\sigma_{NCR}$ and βlobe/Si1 conformations in apo-RNAP (middle), RPc (left) and R_DksA/ppGpp (right, DNA is removed from the RP1_DksA/ppGpp). RNAP (subunits and domains), DksA and DNA are indicated. Close-up views of the RNAP cleft are shown below. The DNA loading gate is closed in the apo-RNAP due to βgate loop (βGL) contacts $\sigma_{1.1}/\sigma_{1.2}$ (white dashed oval). The opening of the DNA loading gate in RPc and RP1_DksA/ppGpp is indicated by blue and red arrows, respectively. **B)** A proposed model of $\sigma_{1.1}$ ejection in the RPc. The RPc is depicted as a transparent surface with cartoon models of the clamp (purple) and lobe/Si1 (blue). The clamp in an apo-form RNAP and the lobe/Si1 in RP1-DksA/ppGpp are colored gray and white, respectively. In the RPc, the σ_{NCR} rotation (back arrow) contacts the β'CT, resulting in clamp movement toward $\sigma_{1.1}$ (red arrow) and a steric clash with $\sigma_{1.1}$ (white oval).

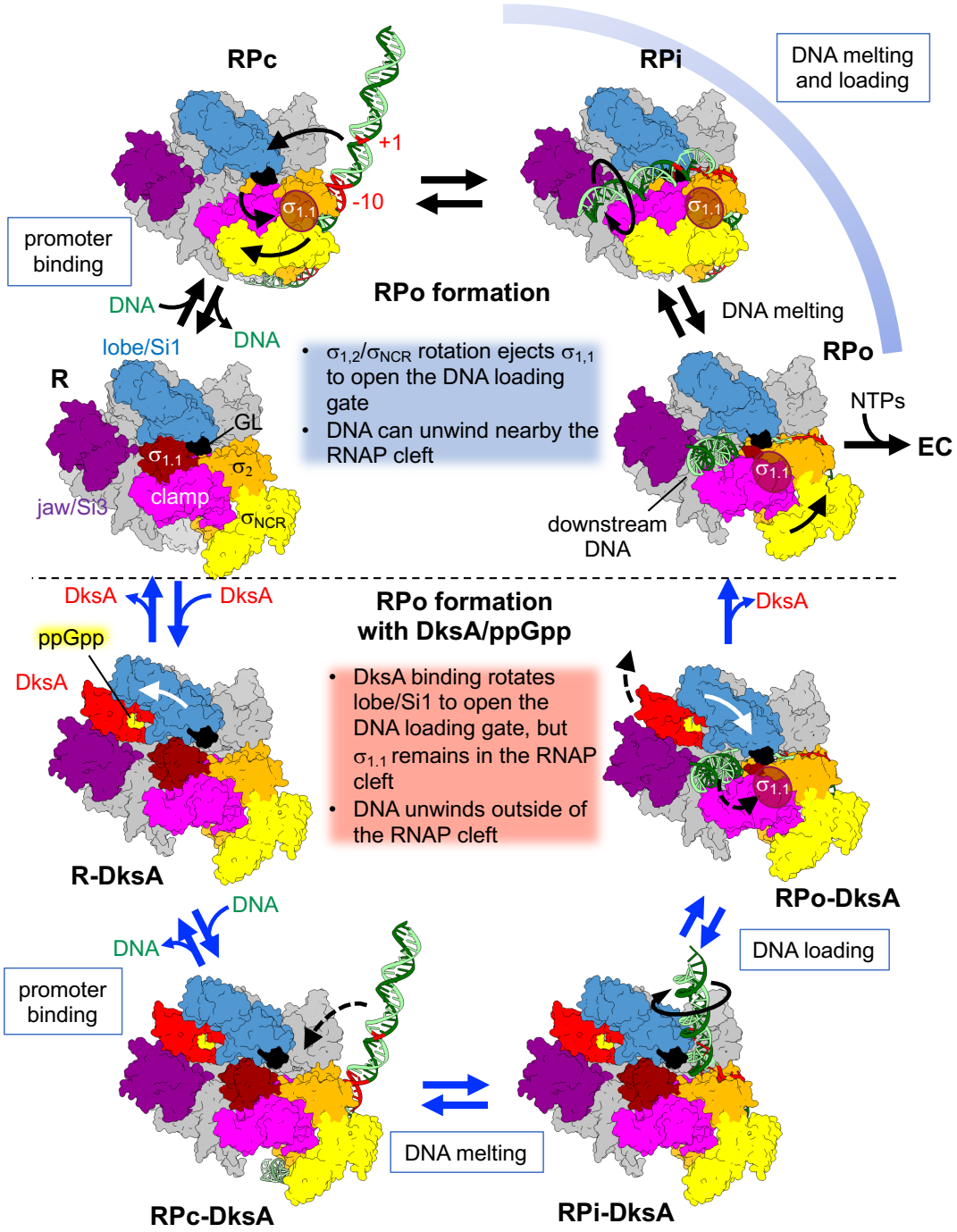


Figure 6. Alternative pathways for open promoter complex formation. Two distinct pathways are shown for open promoter complex formation without (top, blue caption) and with DksA/ppGpp (bottom, red caption). The RNAP holoenzyme (clamp, pink; σ_{NCR}, yellow; $\sigma_{1.1}$, brown; σ_2, orange; βlobe/Si1, light blue; jaw/Si3, purple; rest of RNAP, gray), promoter DNA (tDNA, dark green; ntDNA, light green), DksA (red), and ppGpp (yellow) are shown. Only RPi is a hypothetical intermediate, but others (RPc, RPo, R-DksA, RPc-DksA, RPi_DksA and RPo-DksA) represent the structures determined in this and previous studies. EC, elongation complex.