1960–1970 *Nucleic Acids Research, 2023, Vol. 51, No. 4* *Published online 30 January 2023*
https://doi.org/10.1093/nar/gkad016

The structural mechanism for transcription activation by *Caulobacter crescentus* GcrA

Xiaoxian Wu[1,†], Chengzhi Yu[1,2,†], Wenhui Mu[3], Zhanxi Gu[1,2], Yu Feng[4] and Yu Zhang[1,*]

[1]Key Laboratory of Synthetic Biology, Center for Excellence in Molecular Plant Sciences, Shanghai Institute of Plant Physiology and Ecology, Chinese Academy of Sciences, Shanghai 200032, China, [2]University of Chinese Academy of Sciences, Beijing 100049, China, [3]Key Laboratory of Plant Stress Biology, State Key Laboratory of Cotton Biology, School of Life Sciences, Henan University, Kaifeng 475004, China and [4]Department of Biophysics, and Department of Infectious Disease of Sir Run Run Shaw Hospital, Zhejiang University School of Medicine, Hangzhou, China

Received September 15, 2022; Revised December 28, 2022; Editorial Decision January 03, 2023; Accepted January 05, 2023

ABSTRACT

Canonical bacterial transcription activators bind to their cognate *cis* elements at the upstream of transcription start site (TSS) in a form of dimer. *Caulobacter crescentus* GcrA, a non-canonical transcription activator, can activate transcription from promoters harboring its *cis* element at the upstream or downstream of TSS in a form of monomer. We determined two cryo-EM structures of *C. crescentus* GcrA-bound transcription activation complexes, GcrA TAC_U and GcrA TAC_D, which comprise GcrA, RNAP, σ^{70} and promoter DNA with GcrA *cis* elements at either the upstream or downstream of TSS at 3.6 and 3.8 Å, respectively. In the GcrA-TAC_U structure, GcrA makes bipartite interactions with both σ^{70} domain 2 (σ^{70}_2) and its *cis* element, while in the GcrA-TAC_D structure, GcrA retains interaction with σ^{70}_2 but loses the interaction with its *cis* element. Our results suggest that GcrA likely forms a functionally specialized GcrA-RNAP-σ^A holoenzyme, in which GcrA first locates its *cis* element and then facilitates RNAP to load on core promoter at its proximal region. The sequence-specific interaction of GcrA and DNA is disrupted either at the stage of RPo formation or promoter escape depending on the location of GcrA *cis* elements relative to TSS.

INTRODUCTION

In bacteria, gene transcription is performed by a single RNA polymerase (RNAP) but tightly regulated by transcription factors, small molecules and sequences embedded in genomic DNA (1–5). During transcription initiation, an array of bacterial RNAP holoenzymes, which are composed of RNAP core enzyme and distinct initiation σ factors, recognize their respective promoter DNA and turn on transcription of the downstream genes (6–8). The transcription output of a specific gene is further tuned by transcription factors that bind their cognate *cis* elements at the proximal region of promoter DNA (3,4).

Most transcription factors function at the transcription initiation stage. They alter the preference of RNAP holoenzyme over promoters and fine-tune RNAP holoenzyme activity by increasing or lowering energetic barriers that RNAP must overcome during the complicated process of transcription initiation (9–21). Decades of genetic, biochemical, and structural studies suggested that, in order to augment transcription output, a canonical Class I/II transcription activator functions as a dimer, binds its *cis* element at the upstream of core promoter region, and establishes interactions with RNAP surface patches (RNAP-α subunit, σR4, etc.) (22–25). The bipartite interaction of transcription factors with promoter DNA and RNAP tethers RNAP and promoter together and also accelerates unwinding of promoter DNA (3). Considering that canonical bacterial transcription factors are typically composed of DNA-binding and signal-sensing domains with limited freedom of inter-domain movement, and that the bipartite interaction in a transcription activation complex involves sequence-specific recognition of promoter DNA as well as surface patch-specific recognition of RNAP, the *cis* element of a transcription activator typically locates at regularly fixed locations of the upstream of the transcription start site (TSS) to allow optimal interactions among RNAP, promoter DNA and the transcription activator (4,18,22,23).

However, GcrA in α-proteobacteria activates transcription through a mechanism distinct from the canonical transcription factors (26,27). GcrA is a master regulator of cell cycle progression (26,28,29). It controls 140 transcription units comprising more than 200 genes in *Caulobacter crescentus* involving in nucleotide synthesis, DNA repair, chromosome organization and segregation, and cell division (26,28,30–32). The small GcrA protein

*To whom correspondence should be addressed. Tel: +86 21 54924351; Email: yzhang@cemps.ac.cn
†The authors wish it to be known that, in their opinion, the first two authors should be regarded as Joint First Authors.

(173 residues) comprises a N-terminal DNA binding domain (GcrA DBD) that recognizes a subset of N^6-adenine methylated (m^6A) GANTC sites (YGAKTCG), and a C-terminal σ^{70}-interacting domain (GcrA SID) that makes interaction with domain 2 of the principal σ factor, σ^{70} (26,27,31) (Figure 1A). The two domains are connected by an unstructured linker of ∼60 residues. Intriguingly, in contrast to the regularly fixed locations of *cis* elements at the upstream of TSS for canonical bacterial transcription activators, GcrA *cis* elements are distributed in varied locations either at the proximal upstream (mostly centered at –20 ± 1 and –30 ± 2 positions in the spacer region between the –35 and –10 elements) or immediate downstream of TSS of its activated genes (ranging from +5 to +30) (26,27).

To understand the unique interaction mode of GcrA with RNAP and its *cis* element, we have previously determined a crystal structure of GcrA SID-σ^{70}_2 binary complex and a crystal structure of GcrA DBD–DNA binary complex. The structures show that GcrA SID contacts a large hydrophobic surface patch on σ^{70}_2, explaining its tight interaction with RNAP, and show that GcrA DBD recognizes its methylated *cis* element through its methyl group- and base-specific interactions. The previous results provide the atomic details of interactions required for the transcription activation by GcrA(27), but it remains unclear how GcrA activates transcription from promoters containing its *cis* element at various locations, especially promoters containing its *cis* elements at either upstream or downstream of TSS. In this study, we report two cryo-EM structures of GcrA transcription activation complexes (GcrA TACs), GcrA TAC_U and TAC_D. The GcrA-TAC_U structure shows that GcrA forms a bipartite interaction with σ^{70} and promoter DNA with GcrA *cis* element at the upstream of TSS at the stage of RPo, while the GcrA-TAC_D structure shows GcrA retains interaction with σ^{70}_2 but loses interaction with its *cis* DNA element at the downstream of TSS at the stage of RPo. We propose that GcrA anchors RNAP through GcrA SID and scans its *cis* elements through its flexible GcrA DBD. The interaction of GcrA with its *cis* element at the proximal region of TSS initially helps engagement of RNAP on core promoter region. The GcrA- *cis* element interaction is subsequently disrupted at different stages of transcription initiation according to the locations of GcrA *cis* elements.

MATERIALS AND METHODS

Plasmids

The plasmid information is summarized in Table S2. pET28a-TEV-*gcrA* (1–173) was constructed by ligation of a Nco1-*gcrA*-EcoR1 DNA fragment (amplified from pET28b-His_6-*gcrA* with primer listed in Table S3) into pET28a using T4 ligase (New England Biolabs). pCOLADuet-*rpoB*-*rpoC*-His_6 was constructed by incorporating into pCOLADuet the Nco1-*rpoB*-HindIII and Nde1-*rpoC*-His_{10}-Kpn1 fragments using a homogenous recombination method. pACYCDuet-*rpoA*-*rpoZ* was constructed by incorporating into pACYCDuet the Nco1-*rpoA*-HindIII and Nde1-*rpoZ*-Kpn1 fragments using a homogenous recombination method (NovoRec Plus One Step PCR Cloning kit; Novoprotein). pEASY-P*csM* was constructed by ligating the *csM* promoter region (–150 to +100) with pEASY-blunt vector (TransGen Biotech). pEASY-*PmipZ* was constructed by ligating the *mipZ* promoter region (–50 to +50) and a tR2 terminator (31 nt) with pEASY-blunt vector. The GcrA boxes (GANTC) were replaced by TTTTT in pEASY-P*csM* (–20M) and pEASY-P*mipZ* (+4M/+12M).

Caulobacter crescentus RNAP core enzyme

The *C. crescentus* RNAP core enzyme was purified from *Escherichia coli* strain BL21(DE3) (Novo protein, Inc.) carrying pCOLA-*CcrpoB*-*CcrpoC*-His_{10} and pACYC-*CcrpoA*-*CcrpoZ*. The protein expression was induced at OD_{600} 0.6 with 0.3 mM IPTG at 18°C for overnight. Cells were harvested and resuspended in lysis buffer [40 mM Tris–HCl pH 7.7, 200 mM NaCl, 5% glycerol, 2 mM EDTA, 2 mM DTT, 0.1 mM phenylmethylsulfonyl fluoride (PMSF) and protease inhibitor cocktail (APExBIO, Inc.)] and then lysed using an Avestin EmulsiFlex- C3 cell disrupter (Avestin, Inc.). The supernatant was precipitated by dropwise addition of 10% polyethylenimine (PEI) to a final concentration of 0.25% at 4°C. The pellet was washed with 150 ml of 10 mM Tris–HCl pH 7.7, 200 mM NaCl, 5% glycerol, 1 mM DTT for three times and then dissolved with 300 ml of 10 mM Tris–HCl pH 7.7, 1 M NaCl, 5% glycerol, 1 mM DTT. The supernatant was precipitated again by ammonium sulfate (final concentration; 30 g/100 ml) and retrieved in 150 ml NTA-binding buffer (10 mM Tris–HCl pH 7.7, 500 mM NaCl, 5% glycerol, 5 mM β-mercaptoethanol). The sample was applied onto a Ni-NTA column (SMART, Inc.), washed, and eluted with Ni-NTA buffer containing 300 mM imidazole. The eluted fractions were diluted and loaded onto a Mono Q column (Mono Q 10/100 GL, Cytiva) followed by a salt gradient of buffer A (10 mM Tris–HCl pH 7.7, 150 mM NaCl, 5% (v/v) glycerol, 1 mM DTT, 0.1 mM EDTA) and buffer B (10 mM Tris–HCl, pH 7.7, 500 mM NaCl, 5% (v/v) glycerol, 1 mM DTT, 0.1 mM EDTA). The RNAP fractions were concentrated to 10 mg/ml and stored at −80°C.

Caulobacter crescentus σ^{70} and GcrA

C. crescentus σ^{70} and GcrA was overexpressed in *E. coli* BL21 (DE3) cells carrying pET28a-TEV-*gcrA* (1–173) *or* pET28b-His_6-*rpoD* and purified as described in (27).

Caulobacter crescentus RNAP holoenzyme

C. crescentus RNAP holoenzyme was reconstituted with *C. crescentus* RNAP core enzyme and σ^{70} in a 1:4 molar ratio and then purified with a Superose 6 10/300 GL column (Cytiva) as described in (27).

Nucleic-acid scaffold for cryo-EM structure determination

Nucleic-acid scaffolds for *C. crescentus* GcrA TAC_U were prepared as follows: the nontemplate-strand DNA (5′-AAAAAAAGAGTTGACACCCGGGCG(m^6A)AT

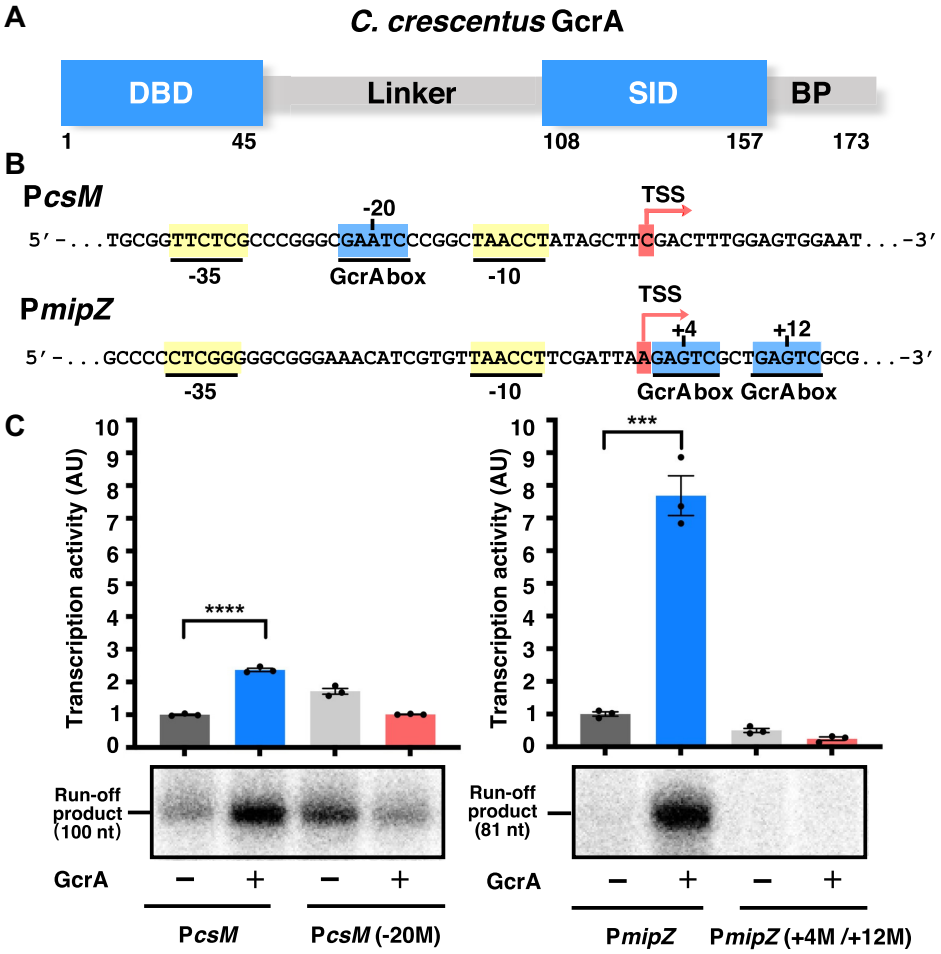

Figure 1. GcrA activates transcription from promoters with GcrA box at either side of TSS. (**A**) The schematic presentation of GcrA domains. The DNA-binding domain (DBD), unstructured linker region (linker), σ factor-interacting domain (SID) and the basic patch region (BP) are labeled. (**B**) The sequence of *C. crescentus csM* and *mipZ* promoter DNA used for multiple-round *in vitro* transcription reactions. Yellow, the –35 and –10 elements; blue, the GcrA boxes; red, transcription start site (TSS). (**C**) GcrA increases RNA output from P*csM* and P*mipZ*. Mutation of GcrA *cis* element abolishes the activity of GcrA. The run-off transcripts were quantified. Data are presented as mean ± SD, $n = 3$ independent experiments. $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$ from two-tailed unpair *t*-tests.

CCCTGCTATAATGGGAGCTGTCACGGATGC-3′), template-strand DNA (5′-GCATCCGTGAGTCGAGG GTAATAAGCAGGG(m^{6}A)TTCGCCCGGGTGTCAA CTCTTTTTTT-3′) and RNA (5′-CUCGA-3′) in annealing buffer (5 mM Tris–HCl, pH 8.0, 200 mM NaCl and 10 mM $MgCl_2$) were heated for 5 min at 95°C and cooled to 22°C in 2°C steps with 30 s per step using a thermal cycler.

Nucleic-acid scaffolds for *C. crescentus* GcrA TAC_D were prepared as follows:

nontemplate-strand DNA (5′-CTGTGGATTGAGGCC CTTAGCCCCTTGACAGGCGGGAAACATCGTGT TAACCTTCGATTAAG(m^{6}A)GTCGCTG(m^{6}A)GTCG CGACGCAGAAGGATC-3′), template-strand DNA (5′-GATCCTTCTGCGTCGCG(m^{6}A)CTCAGCG(m^{6}A)CT CTTAATCGAAGGTTAACACGATGTTTCCCGCC TGTCAAGGGGCTAAGGGCCTCAATCCACAG-3′) in annealing buffer (5 mM Tris–HCl, pH 8.0, 200 mM NaCl and 10 mM $MgCl_2$) were heated for 5 min at 95°C and cooled to 22°C in 2°C steps with 30 s per step using a thermal cycler.

Cryo-EM sample preparation: *C. crescentus* GcrA TAC_U and GcrA TAC_D

The *C. crescentus* GcrA-TAC_U complex was reconstituted by mixing *C. crescentus* RNAP holoenzyme, GcrA, and the nucleic-acid scaffold in a 1:4:1.3 molar ratio. The mixture was applied to a Superose 6 10/300 GL column (Cytiva) equilibrated in 10 mM HEPES, pH 7.5, 150 mM KCl, 5 mM $MgCl_2$ and 5 mM DTT. Fractions containing *C. crescentus* GcrA TAC_U were collected and concentrated to ∼12 mg/ml. The freshly purified *C. crescentus* GcrA-TAC_U at 12 mg/ml was incubated with 3-([3-cholamidopropyl] dimethylammonio)-2-hydroxy-1-propanesulfonate (CHAPSO, 8 mM final; Hampton Research) before grid preparation. About 3 μl of *C. crescentus* GcrA TAC_U was applied on a glow-discharged C-flat CF-1.2/1.3 400 mesh holey carbon grid (Protochips), blotted with Vitrobot Mark IV (FEI), and plunge-frozen into liquid ethane with 95% chamber humidity at 9.5°C. To prepare *C. crescentus* GcrA TAC_D complex, the nucleic-acid scaffold was incubated with RNAP

for 25s and the mixture was immediately applied on a glow-discharged C-flat CF-1.2/1.3 400 mesh holey carbon grid (Protochips), blotted with Vitrobot Mark IV (FEI), and plunge-frozen into liquid ethane with 95% chamber humidity at 9.5°C.

Cryo-EM data collection and processing: *C. crescentus* GcrA TAC_U

The *C. crescentus* GcrA-TAC_U dataset was collected on a FEI Titan Krios equipped at 300 keV with a K2 Summit direct electron detector (Gatan). The images were recorded using Serial EM software at a nominal magnification of 22 500 (1.014 Å/pixel) in counting mode. Each image movie of 32 frames was collected by exposure of 8.0 s to give a total electron exposure of 62.4 electrons/Å2 (flux: 8 electrons/pixel/s). The 3626 images were collected with defocus range from −1.2 to −2.2 μm. Frames of individual movies were aligned using MotionCor2, and contrast-transfer-function estimations were performed using CTFFIND4. Image processing was performed with RELION 3.0. A total of 481 494 particles were auto picked and extracted from the dataset by using the 2D classified templates generated from previous *E. coli* RNA polymerase elongation complex. The particles were subjected to 2D classification ($N = 50$, iterations 25) and the good 2D classes were subjected to 3D classification ($N = 4$, iterations = 25) using a 40 Å low-pass-filtered cryo-EM structure of *E. coli* RNA polymerase elongation complex as the initial model. The 3D class, which contains 309 114 particles and shows clear feature of *C. crescentus* RNA polymerase, GcrA and promoter DNA were selected and refined to 4.16 Å. The particles (309 114) were further subjected to focused 3D classification ($N = 2$, iterations = 25, without alignment) by subtracting signal outside the black mask. The 287 227 particles from the best-resolved class were reverted and subjected to auto refinement, CTF refinement, particle polishing, a second round of auto refinement, and postprocess, resulting in the final map. The Gold-standard Fourier-shell-correlation analysis indicated a nominal resolution of 3.62 Å at 0.143 FSC cutoff.

The crystal structures of GcrA-SID/σ^{70}_2 (PDB: 5YIX) and GcrA-DBD/DNA (PDB: 5YIV) and cryo-EM structure of *E. coli* RPo (PDB:7MKL) were fit into the cryo-EM map (27,33). The iterative cycles of model building in Coot (Ramachandran, trans peptide, planar peptide restraints applied) and refinement in Phenix were performed.

Cryo-EM data collection and processing: *C. crescentus* GcrA TAC_D

The *C. crescentus* GcrA-TAC_D dataset was collected on a FEI Titan Krios equipped at 300 keV with a K2 Summit direct electron detector (Gatan). The images were recorded using Serial EM software at a nominal magnification of 22 500 (1.0 Å/ pixel) in counting mode. Each image movie of 32 frames was collected by exposure of 7.6 s to give a total electron exposure of 60.8 electrons/$Å^2$ (flux: 8 electrons/pixel/s). The 1886 images were collected with defocus range from −1.2 to −2.2 μm. Frames of individual movies were aligned using MotionCor2, and contrast-transfer-function estimations were performed using CTFFIND4. Image processing was performed with RELION 3.0. A total of 865 129 particles were auto picked and extracted from the dataset by using the 2D classified templates generated from previous *C. crescentus* GcrA TAC_U. The particles were subjected to 2D classification ($N = 50$, iterations 25) and the good 2D classes were subjected to 3D classification ($N = 4$, iterations = 25) using a 40 Å low-pass-filtered cryo-EM structure of *C. crescentus* GcrA TAC_U as the initial model. The 3D class, which contains 109 869 particles and shows clear feature of *C. crescentus* RNA polymerase, GcrA and promoter DNA were selected and subjected to auto refinement, CTF refinement, particle polishing, a second round of auto refinement, and postprocess, resulting in the final map. The Gold-standard Fourier-shell-correlation analysis indicated a nominal resolution of 3.79 Å at 0.143 FSC cutoff.

The structure of *C. crescentus* GcrA TAC_U was fit into the cryo-EM map. The iterative cycles of model building in Coot (Ramachandran, trans peptide, planar peptide restraints applied) and refinement in Phenix were performed.

In vitro transcription assays

Template DNAs for *in vitro* transcription assay were amplified by PCR using plasmids pEASY-P*csM*, pEASY-P*csM* (−20M), pEASY-P*mipZ* or pEASY-P*mipZ* (+4M/+12M) as templates. Reactions were performed in transcription buffer (40 mM Tris–HCl pH 8.0, 75 mM NaCl, 5 mM $MgCl_2$, 2.5 mM DTT, 12.5% glycerol, 50 μg/ml BSA). The reaction mixtures (20 μl) containing 40 nM *C. crescentus* RNAP holoenzyme, 640 nM (final concentrations) GcrA or GcrA (ΔBP) were incubated at 37 °C for 5 min. The reactions were initiated by addition of 1.2 μl NTP-promoter mixture containing 100 μM ATP, 100 μM UTP, 100 μM GTP, 100 μM CTP, 0.165 μM [α-^{32}P] UTP and 40 nM promoter DNA (final concentrations) and allowed to proceed for 20 min at 37 °C, and then stopped by addition of 5 μl stop buffer (8 M urea, 20 mM EDTA, 0.025% xylene cyanol and 0.025% bromophenol blue). The sample were boiled at 95°C for 5 min, immediately cooled in ice for 5 min, and electrophoresed on 15% (19:1 acrylamide/bisacrylamide) urea–polyacrylamide slab gels. The radiograph was obtained by storage-phosphor scanning (Typhoon; GE Healthcare Inc.).

Fluorescence polarization

GcrA was labeled with fluorescein on an exogenous cysteine at the N-terminus of GcrA. The labeling reaction mixture (2 ml) containing GcrA (0.07 mM) and fluorescein-5-maleimide (0.7 mM; ThermoFisher Scientific, Inc.) in 10 mM Tris–HCl, pH 7.7, 100 mM NaCl, 1% glycerol was incubated overnight at 4°C. The reaction was terminated by addition of 2 μl DTT (1 M; stock concentration), and the reaction mixture was loaded onto a 10 ml PD-10 desalting column (GE healthcare, Inc.). The fractions containing labeled protein was pooled and concentrated to 3 mg/ml.

Cy5-labeled *PmipZ* fragment was amplified from pEASY-*PmipZ* with a pair of primers 5′-Cy5-GATCCTTCTGCGTCGCGACT-3′ and

5′-CTGTGGATTGAGGCCCTTAGCC-3′, and then methylated and purified as described in (26).

To measure the binding affinity of GcrA to RNAP, the fluorescein-labeled GcrA (8 nM) was incubated with *Cc* RNAP holoenzyme (8, 16, 32, 64, 128, 256, 512 or 1024 nM). To measure the binding affinity of GcrA to promoter DNA, the cy5-labeled DNA (4 nM) was incubated with GcrA (1, 2, 4, 8, 16, 32, 64 or128 μM). To measure the binding affinity of RNAP to promoter DNA, the cy5-labeled DNA (4 nM) was incubated with *Cc* RNAP holoenzyme (4, 8, 16, 32, 64, 128, 256 or 512 nM). The reaction mixtures were incubated in 100 μl FP buffer (10 mM Tris–HCl, pH 7.7, 100 mM NaCl,1 mM DTT, 1% glycerol and 0.025% Tween-20) in a 96-well plate (Corning, Inc) for 5 min at room temperature. The fluorescence polarization (FP) signals were measured using a plate reader (SPARK, TECAN Inc.; excitation and emission filters of 485/20 and 520/20 nm, respectively for fluorescein; excitation and emission filters of 635/35 and 665/8 nm, respectively for Cy5). The data were plotted in GraphPad Prism version 8.4.0 (GraphPad software, Inc.) and the dissociation constant Kd were estimated by fitting the data to the following equation,

$$F = B[S]/(K_d + [S]) + F_0$$

where F is the FP signal at a given concentration of RNAP, F_0 is the FP signal in the absence of RNAP, $[S]$ is the concentration of RNAP, and B is an unconstrained constant.

RESULTS

GcrA activates transcription of RNAP from two types of promoter DNA

To verify that GcrA can activate transcription of RNAP from promoter DNA with GcrA *cis* element (GcrA box) at either the upstream or downstream of TSS, we measured the effect of GcrA on RNAP transcription from two representative promoter DNAs (26), P*csM* containing one predicted GcrA box at the upstream of TSS centered at −20 position, and P*mipZ* containing two predicted GcrA boxes at the downstream of TSS centered at positions +4 and +12 (Figure 1B). Our *in vitro* transcription results show that GcrA substantially increases RNA outputs from both promoter DNAs (Figure 1C), confirming that GcrA can activate transcription of promoters with its *cis* element at either the upstream or downstream of TSS. Moreover, mutating the GcrA boxes in either P*csM* or P*mipZ* abolishes GcrA activity (Figure 1C), suggesting the critical role of GcrA-*cis* DNA interaction in the transcription activation activity of GcrA. Most GcrA-activated promoters contain only one GcrA box (26). We suspect that only one of the two boxes in P*mipZ* plays a major role, most likely the one centered at +12 position (TGAGTCG), as its sequence is closer to the consensus sequence (YGAKTCG) of GcrA-recognized motif. In short, our results provide direct evidence to support the notion that GcrA can activate promoters with its *cis* elements at either upstream or downstream of TSS (26).

The cryo-EM structure of *C.crescentus* GcrA TAC_U

To understand the structural basis underlying the transcription activation of GcrA on both types of promoters, we sought to determine transcription activation complexes of GcrA comprising *C. crescentus* RNAP holoenzyme, GcrA, and two promoter DNAs (P*csM* for GcrA TAC_U; P*mipZ* for GcrA TAC_D). Our previous work indicates that GcrA likely retains interactions with RNAP and promoter DNA during the process of RPo formation if GcrA *cis* element is in the spacer region between the −35 and −10 elements (27). Therefore, we assembled GcrA-TAC_U complex using *C. crescentus* GcrA, RNAP-σ^{70} holoenzyme, and a pre-melted derivative of P*csM* (Figures S1A and B) that comprises an upstream DNA duplex (−45 to −12, respective to TSS) with a methylated GcrA *cis* element (Gm^6AATC, -22 to -18) and the consensus UP and -35 elements, a pre-melted transcription bubble (−11 to +2) with the consensus −10 and discriminator elements, a 5-nt RNA primer complementary to template DNA (−4 to +1), and a downstream DNA duplex (+3 to +12) (Figure 2A). The complex structure was determined through a cryo-EM method (Table S1).

The cryo-EM map of GcrA TAC_U reconstructed to a resolution of 3.6 Å (Figures S1C–G) shows clear signals for *C. crescentus* RNAP holoenzyme and promoter DNA (−45 to + 12). The map also shows clear signals for a monomeric GcrA SID (residues 105−155) associated with $\sigma^{70}{}_2$ and a monomeric GcrA DBD (residues 1 to 45) associated with its *cis* element at the −35/−10 spacer of promoter DNA (Figures 2B, C and 3A), confirming that GcrA functions as a monomer (27). Intriguingly, the cryo-EM map shows additional signals for part of the GcrA linker region that travels underneath $\sigma^{70}{}_{NCR}$ following the minor groove of the promoter DNA (Figure 2B). In summary, the structure suggests that GcrA functions as a monomer, SID and DBD of GcrA interact with $\sigma^{70}{}_2$ and promoter DNA respectively and independently, and the linker travels on the front face of RNAP to connect GcrA DBD and SID (Figures 2B and C).

The GcrA-TAC_U structure is superimposable on the crystal structure of GcrA-DBD/DNA as well as the crystal structure of GcrA-SID/$\sigma^{70}{}_2$ (Figures 3B and C) (27). Notably, the cryo-EM map of GcrA-TAC_U structure reveals additional interactions made by the C-terminal basic patch of GcrA SID with the non-template single-stranded DNA of the transcription bubble. The weak but continuous map, which extends from the C terminus of the crystal structure model of GcrA SID, could be assigned to 4 residues (157–160) of the GcrA C-terminal basic patch. Three conserved positively charged arginine residues (K157, K158 and K159) in the basic batch are in close proximity to and could make potential interactions with the phosphate backbones of discriminator nucleotides (−7 to −4) of the non-template strand (Figures 2D, 3A and D). The interactions are expected to restrain the non-template ssDNA of the transcription bubble in the main cleft and thus increase RPo stability (Figures 2D and 3D). In agreement with the structure, removal of the basic patch reduced the transcription activation activity of GcrA in an *in vitro* transcription assay (Figure 3E). Notably, GcrA (ΔBP) remains partial activation activity likely due to the interaction between DNA and GcrA DBD (Figure 3E). The interaction between GcrA basic patch and discriminator element also explains the residual activation activity of a DBD-

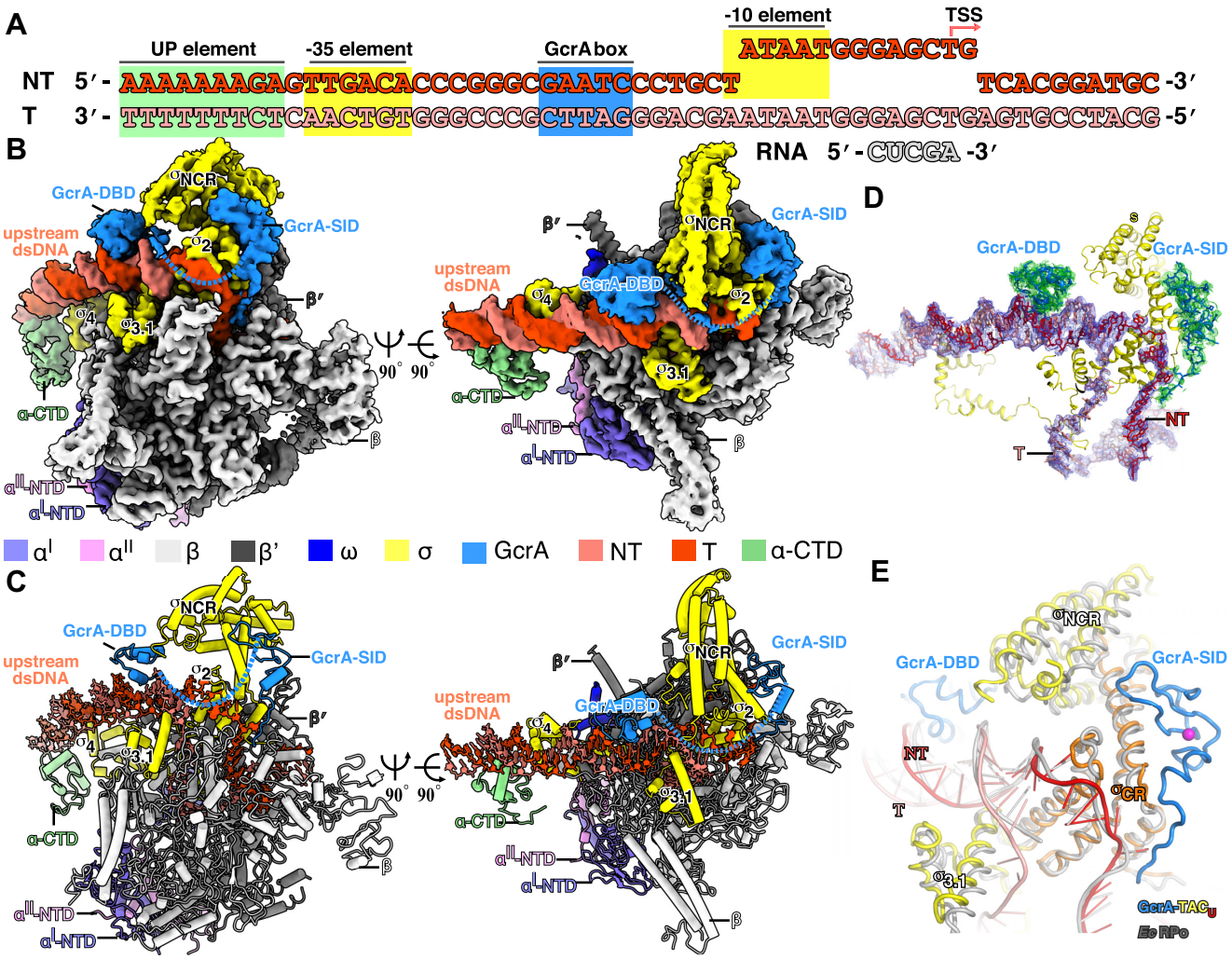

Figure 2. The cryo-EM structure of *C. crescentus* GcrA TAC_U. (**A**) The nucleic-acid scaffold used for structure determination of *C. crescentus* GcrA TAC_U. Red, the nontemplate DNA; pink, template DNA; gray, RNA; red arrow, TSS. (**B**) The side and top views of the cryo-EM map of *C. crescentus* GcrA TAC_U. (**C**) The side and top views of the overall structure of *C. crescentus* GcrA TAC_U. RNAP subunits, GcrA and nucleic-acid scaffold are colored as the color scheme. (**D**) The cryo-EM map and structure model of the nucleic-acid scaffold (blue mesh) and GcrA (green mesh). (**E**) The structure superimposition of *C. crescentus* GcrA TAC_U and *E. coli* RPo (PDB: 6CA0).

truncated GcrA derivative (27). In short, we show that the C-terminal basic patch of GcrA makes interaction with the non-template ssDNA and the interaction might stabilize RPo.

It is worth noting that the σ^{70}_{NCR} in the GcrA-TAC_U structure adopts a different conformation compared to that of our previously reported crystal structure of *C. crescentus* σ^{70}_2/GcrA SID (Figures 4A and B). In the cryo-EM structure of GcrA TAC_U, σ^{70}_{NCR} sits on top of the upstream dsDNA and is loosely sandwiched between GcrA DBD and GcrA SID (Figure 3A). However, in our previous crystal structure of σ^{70}_2/GcrA SID, the σ^{70}_{NCR} rotates ~20° in anti-clockwise direction along an axis perpendicular to the coiled-coil plane compared with that in the GcrA-TAC_U structure. Superimposition of the two structures shows that σ^{70}_{NCR} prevents dsDNA binding in the crystal structure of σ^{70}_2/GcrA SID, suggesting it is in an inactive conformation (Figure 4C). We infer that such inactive conformation of σ^{70}_{NCR} might also exist in the context of RNAP-σ^{70} holoenzyme to prevent non-specific dsDNA binding and to allow engagement of genuine promoters. Intriguingly, *E. coli* σ^{70}_{NCR} adopts another conformation in recently reported cryo-EM structures of *E. coli* transcription initiation complexes, in which σ^{70}_{NCR} rotates in clockwise direction along the axis of the coiled-coil to a position, where σ^{70}_{NCR} contacts β' clamp head domain to facilitate either ejection of $\sigma^{70}_{1.1}$ or promoter escape (13,17). Our structure and previous results together highlight conformational flexibility of σ^{70}_{NCR} accounting for its multiple function in different stages of transcription initiation.

The cryo-EM structure of *C. crescentus* GcrA TAC_D

To understand how GcrA activates promoter DNA with GcrA *cis* element at the proximal downstream of TSS, we incubated the GcrA-bound RNAP with a P*mipZ* derivative for 25 seconds and vitrified the reaction mixture (Figures 5A, S2A and B). The sample was subject to data collection

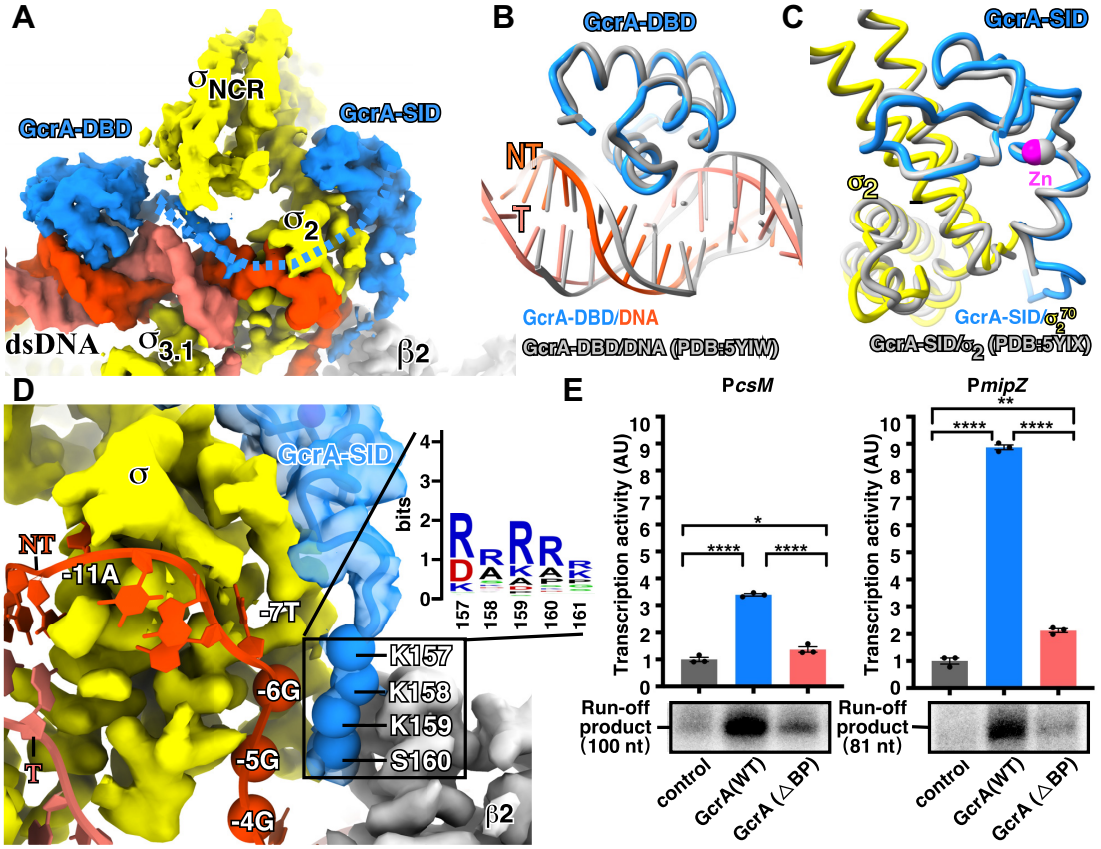

Figure 3. Both GcrA DBD and GcrA SID contribute to the transcription activation activity of GcrA. (**A**) The interaction among GcrA, σ^{70}, nucleic-acid scaffold and RNAP. colors are as above. (**B**) Structure superimposition of GcrA TAC_U (colored as above) and the crystal structure of GcrA-DBD/DNA (gray). (**C**) Structure superimposition of GcrA TAC_U (colored as above) and the crystal structure of GcrA-SID/$\sigma^{70}{}_2$ (PDB: 5YIX; gray). (**D**) The potential interaction between residues of GcrA BP and the backbone phosphates of the discriminator element nucleotides. The insert shows the consensus sequence of GcrA BP. (**E**) Deletion of GcrA BP impairs transcription activation activity of GcrA on both P*csM* and P*mipZ*. GcrA (ΔBP), a GcrA derivative with deleted basic patch (residues 157–163). The run-off transcripts were quantified. Data are presented as mean ± SD, $n = 3$ independent experiments. $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$ from two-tailed unpair *t*-tests.

on a Krios 300 kV transmission electron microscopy. Iterative cycles of 3D classification of single particles resulted in only one major group, from which a cryo-EM map was reconstructed to a resolution of 3.8 Å (Figures S2C–G and Table S1). The map shows that ∼12 bp of the promoter were unwound and the template ssDNA was loaded into the active-site cleft, suggesting we have obtained an RPo complex (Figure 5D). The map also shows clear signal for GcrA SID associating with $\sigma^{70}{}_2$ (Figure 5B). We thereby named the complex as GcrA TAC_D. Structure superimposition of GcrA TAC_U and TAC_D shows GcrA SID makes the same interactions with σ^{70} (Figure 5E).

In the structure of GcrA TAC_D, the downstream promoter DNA is loaded into the downstream DNA channel and fully protected by RNAP (Figure 5D). The two GcrA *cis* elements (centered at +4 and +12, respectively) are buried inside the RNAP main cleft (Figure 5D), explaining the absence of signals for GcrA DBD in the cryo-EM map of GcrA TAC_D. The structure indicates that the GcrA-DNA interaction is disrupted during the progress of RPo formation. The structure also predicts that GcrA DBD–DNA interaction should be weak enough so that limited hindrance is posed in the progress of promoter unwinding. To test the hypothesis, we measured the binding affinity of RNAP and GcrA towards promoter DNA P*mipZ*. The result shows that GcrA binds to P*mipZ* with a Kd value of 48 μM, four orders of magnitude lower than the affinity between RNAP and promoter DNA (5.3 nM; Figure 5F), supporting the hypothesis. Moreover, GcrA binds RNAP with a Kd value of 60 nM, three orders of magnitude lower than that of GcrA and promoter DNA (Figure 5F), consistent with the GcrA-TAC_D structure showing that GcrA SID remains tightly associated with RNAP while GcrA DBD has been displaced and dissociated from the promoter DNA (Figure 5B and C).

DISCUSSION

In summary, we have determined cryo-EM structures of *C. crescentus* GcrA transcription activation complexes (GcrA TAC_U and TAC_D) that comprise two types of promoter DNA with GcrA *cis* element at either the upstream or downstream of TSS. The two cryo-EM structures provide direct evidence that support three unique features of transcription activation by GcrA. First, the structures show that GcrA functions as a monomer and interacts with RNAP in a 1:1 stoichiometric ratio, unlike canonical bacterial transcription activators forming homodimers

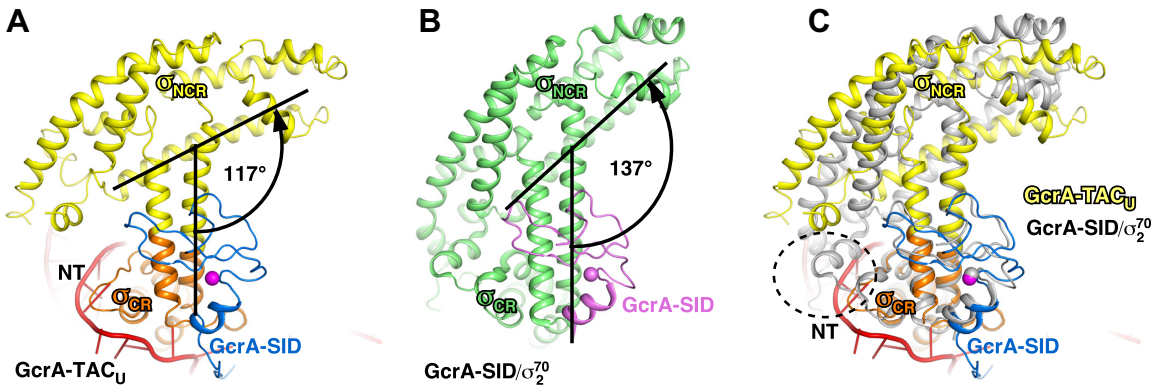

Figure 4. The conformational flexibility of σ_{NCR}. (**A**) The σ_{NCR} conformation in *C. crescentus* GcrA TAC_U. (**B**) The σ_{NCR} conformation in *C. crescentus* GcrA-SID/$\sigma^{70}{}_2$ (PDB:5YIX). (**C**) The structure superimposition of GcrA TAC_U and GcrA-SID/$\sigma^{70}{}_2$ (PDB:5YIX) shows potential clash between σ_{NCR} and the promoter DNA.

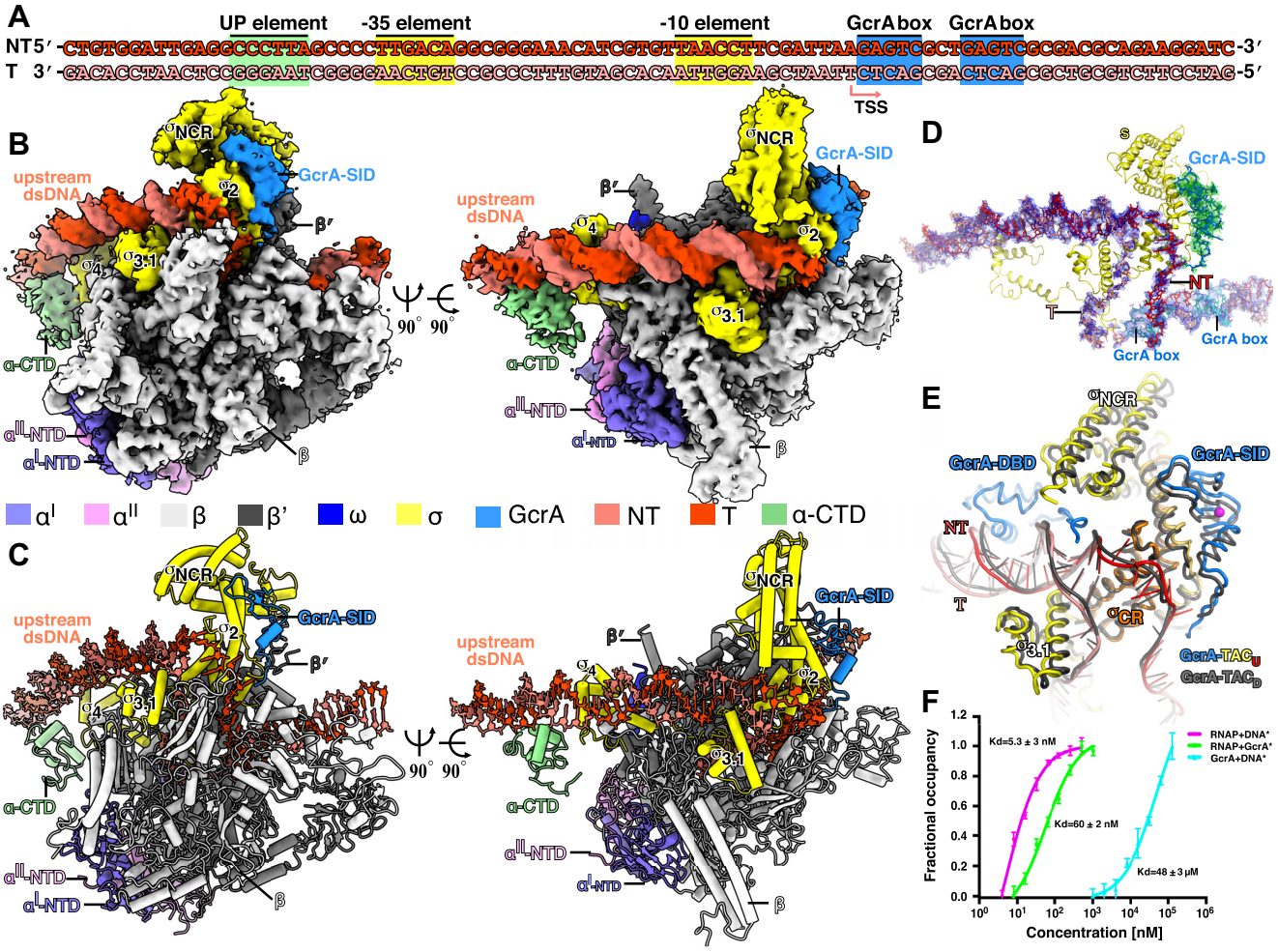

Figure 5. The cryo-EM structure of *C. crescentus* GcrA TAC_D. (**A**) the nucleic-acid scaffold used for *C. crescentus* GcrA TAC_D. (**B**) The front and top views of cryo-EM map of *C. crescentus* GcrA TAC_D. (**C**) The front and top views of the overall structure of *C. crescentus* GcrA TAC_D. RNAP subunits, GcrA and nucleic-acid scaffold are colored as the color scheme. (**D**) The cryo-EM map and structure model of the nucleic-acid scaffold (blue mesh) and GcrA (green mesh). (**E**) The structure superimposition of DNA between *C. crescentus* GcrA TAC_U and GcrA TAC_D. (**F**) The binding affinity among RNAP, GcrA, and the promoter, P*mipZ*. Data are presented as mean ± SD, $n = 3$ independent experiments.

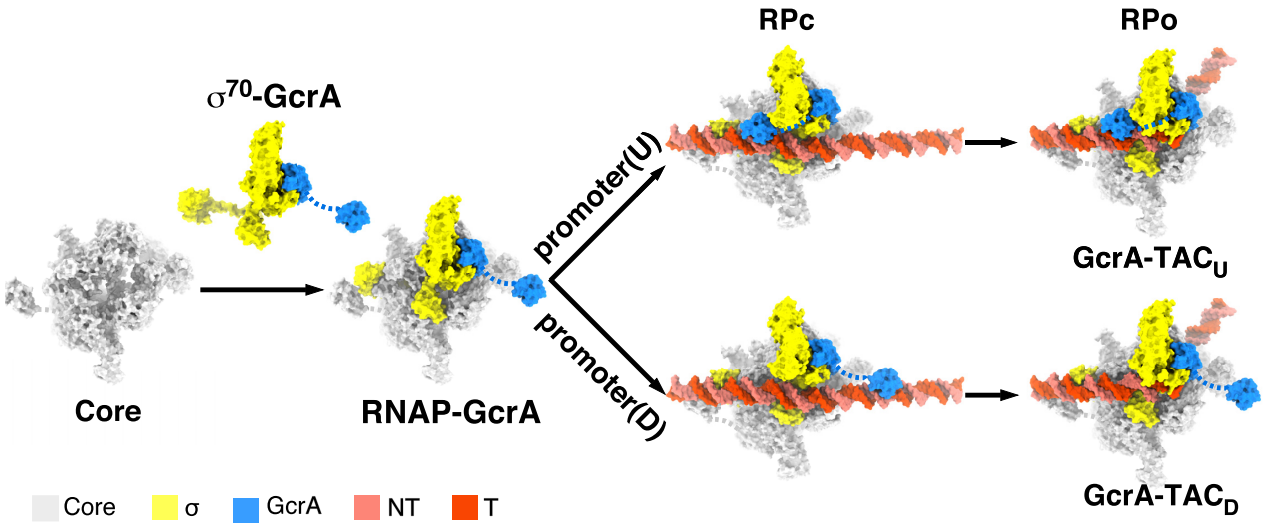

Figure 6. The model of transcription activation by GcrA. RNAP forms a stable complex with GcrA. The long flexible inter-domain linker of GcrA assists the RNAP-GcrA complex to locate the GcrA regulated promoters that comprise GcrA *cis* elements at either the upstream or downstream of TSS. GcrA DBD is either displaced during RPo formation as observed in GcrA TAC_D or retained as observed in GcrA TAC_U. The retained GcrA DBD is likely displaced during promoter escape.

(18–20,22,23,34–36). Second, the structures show that GcrA uses two independent units (SID and DBD of GcrA) to recognize the major subtype of cellular RNAP holoenzyme (RNAP-σ^{70} holoenzyme) and promoters containing GcrA *cis* elements. The two units of GcrA are connected by a long flexible linker and do not communicate with each other, different from canonical bacterial transcription factors that usually transmit signals from the effector domain to the DNA-binding domain (1,18–20,22,23,35,36). Third, the long linker connecting the two independent units confers GcrA large flexibility and allows GcrA DBD to establish interaction with its *cis* elements at a broad range of locations around TSS.

GcrA exhibits high affinity with σ^{70} through a large hydrophobic interface. It is likely the GcrA forms a stable complex with a sub-population of σ^{70}. The GcrA-σ^{70} binary complex forms a functionally specialized RNAP-σ^{70} holoenzyme selectively responsible for transcription of >200 genes directly regulated by GcrA (26). Due to the high affinity between GcrA and σ^{70}, the GcrA-σ^{70} binary complex likely also functions as a unit in the process of transcription initiation, i.e. promoter recognition, promoter unwinding, and promoter escape (Figure 6).

GcrA exhibits significant different affinities towards RNAP and promoter DNA. We propose that such affinity difference accounts for the ability of GcrA to activate transcription from promoters with its *cis* element at either upstream or downstream of TSS. In both structures of GcrA TACs, GcrA SID remains attached on σ^{70}, consistent with the high affinity between GcrA and RNAP-σ^{70} holoenzyme. GcrA makes much weaker interaction with promoter DNA compared with that with RNAP. The interaction is likely sufficient for enriching GcrA-bound RNAP on GcrA-regulated promoters and ensures GcrA DBD is displaced from promoter DNA without significant energetic barrier either during RPo formation as shown in the cryo-EM structure of GcrA TAC_D or during promoter escape as in the case of GcrA TAC_U.

Our model of GcrA activation also explains the unique locations of GcrA *cis* elements on its-activated promoters (26). The GcrA-TAC_U structure shows that GcrA DBD resides in the major groove of its *cis* DNA element located between the −35 and −10 elements. Such interaction can only be established when GcrA DBD and RNAP core enzyme locate at the opposite sides of promoter dsDNA to avoid steric hinderance, explaining that GcrA *cis* elements are centered at –20 ± 1 and –30 ± 2 positions when they are present at the upstream of TSS (26). The GcrA-TAC_D structure predicts GcrA DBD makes interactions with its *cis* element through the stage of RPc, in which the downstream dsDNA remains outside of the DNA main cleft of RNAP and thereby poses little hinderance for GcrA binding, in agreement with the findings that GcrA *cis* elements locate at a broad region downstream of TSS ranging from +5 to +30 (26). The absence of GcrA *cis* element further upstream and downstream is likely due to fixed length of the GcrA DBD/SID linker. The majority of GcrA-activated promoters contain only one GcrA *cis* element (26), consistent with the working model of GcrA. However, the presence of two nearby GcrA boxes might facilitate GcrA activation, because GcrA-RNAP-σ^{70} holoenzyme might have higher affinity with and thus better chance to locate the promoters with two GcrA *cis* elements than the promoters with a single GcrA box in a three-dimensional diffusion model of promoter search (37–39).

It is worth noting that the strategy that GcrA-RNAP-σ^{70} holoenzyme employs to recognize downstream promoter DNA (P*mipZ* as an example) in bacteria resembles the strategy that TFIID-Pol II employs to recognize a subset promoter DNA in human. TFIID employs mobile units on TAF1/7 subunits to recognize the Inr and DPE elements located at the TSS downstream region of promoter DNA.

The interactions are initially established at the stage of promoter recognition but are disrupted upon assembly of the transcription pre-initiation complex (40). GcrA-RNAP-σ^{70} holoenzyme uses GcrA DBD to recognize its *cis* element located at the TSS downstream region of promoter DNA. The interactions are also initially established at the state of RPc but are disrupted upon RPo formation (Figure 6).

In summary, we have provided the detailed structural evidence supporting the unique characteristic of the bacterial transcription factor GcrA. Our work expands the current models of bacterial transcription activation. Our work also provides another example that transcription factors in prokaryotes and eukaryotes employ similar mechanism to regulate transcription at the transcription initiation stage.

DATA AVAILABILITY

The cryo-EM maps and coordinates were deposited in Protein Data Bank and Electron Microscopy Data Bank (*C. crescentus* GcrA TAC_U: 7YE1 and EMD-333761; *C. crescentus* GcrA TAC_D: 7YE2 and EMD-33762).

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Dr Liangliang Kong, Dr Fangfang Wang, Dr Guangyi Li and Dr Jialin Duan at the cryo-EM center of NFPS in Shanghai, Dr Shenghai Chang at the cryo-EM center of Zhejiang University for assistance with data collection.

FUNDING

National Key Research and Development Program of China [2018YFA0900701]; National Natural Science Foundation of China [31870047]. Funding for open access charge: National Natural Science Foundation of China [31870047].
Conflict of interest disclosure. None declared.

REFERENCES

1. Browning,D.F., Butala,M. and Busby,S.J.W. (2019) Bacterial transcription factors: regulation by pick "N" mix. *J. Mol. Biol.*, **431**, 4067–4077.
2. Gourse,R.L., Chen,A.Y., Gopalkrishnan,S., Sanchez-Vazquez,P., Myers,A. and Ross,W. (2018) Transcriptional responses to ppGpp and DksA. *Annu. Rev. Microbiol.*, **72**, 163–184.
3. Chen,J., Boyaci,H. and Campbell,E.A. (2021) Diverse and unified mechanisms of transcription initiation in bacteria. *Nat. Rev. Microbiol.*, **19**, 95–109.
4. Browning,D.F. and Busby,S.J. (2016) Local and global regulation of transcription initiation in bacteria. *Nat. Rev. Microbiol.*, **14**, 638–650.
5. Landick,R. (2021) Transcriptional pausing as a mediator of bacterial gene regulation. *Annu. Rev. Microbiol.*, **75**, 291–314.
6. Feklistov,A., Sharon,B.D., Darst,S.A. and Gross,C.A. (2014) Bacterial sigma factors: a historical, structural, and genomic perspective. *Annu. Rev. Microbiol.*, **68**, 357–376.
7. Sineva,E., Savkina,M. and Ades,S.E. (2017) Themes and variations in gene regulation by extracytoplasmic function (ECF) sigma factors. *Curr. Opin. Microbiol.*, **36**, 128–137.
8. Danson,A.E., Jovanovic,M., Buck,M. and Zhang,X. (2019) Mechanisms of sigma(54)-dependent transcription initiation and regulation. *J. Mol. Biol.*, **431**, 3960–3974.
9. Chen,J., Chiu,C., Gopalkrishnan,S., Chen,A.Y., Olinares,P.D.B., Saecker,R.M., Winkelman,J.T., Maloney,M.F., Chait,B.T., Ross,W. *et al.* (2020) Stepwise promoter melting by bacterial RNA polymerase. *Mol. Cell*, **78**, 275–288.
10. Jensen,D., Manzano,A.R., Rammohan,J., Stallings,C.L. and Galburt,E.A. (2019) CarD and RbpA modify the kinetics of initial transcription and slow promoter escape of the *Mycobacterium tuberculosis* RNA polymerase. *Nucleic Acids Res.*, **47**, 6685–6698.
11. Xu,J., Cui,K., Shen,L., Shi,J., Li,L., You,L., Fang,C., Zhao,G., Feng,Y., Yang,B. *et al.* (2019) Crl activates transcription by stabilizing active conformation of the master stress transcription initiation factor. *Elife*, **8**, e50928.
12. Cartagena,A.J., Banta,A.B., Sathyan,N., Ross,W., Gourse,R.L., Campbell,E.A. and Darst,S.A. (2019) Structural basis for transcription activation by Crl through tethering of sigma(S) and RNA polymerase. *Proc. Natl. Acad. Sci. U.S.A.*, **116**, 18923–18927.
13. Shin,Y., Qayyum,M.Z., Pupov,D., Esyunina,D., Kulbachinskiy,A. and Murakami,K.S. (2021) Structural basis of ribosomal RNA transcription regulation. *Nat. Commun.*, **12**, 528.
14. Chen,J., Gopalkrishnan,S., Chiu,C., Chen,A.Y., Campbell,E.A., Gourse,R.L., Ross,W. and Darst,S.A. (2019) E. coli TraR allosterically regulates transcription initiation by altering RNA polymerase conformation. *Elife*, **8**. e49375.
15. Muller,A.U., Kummer,E., Schilling,C.M., Ban,N. and Weber-Ban,E. (2021) Transcriptional control of mycobacterial DNA damage response by sigma adaptation. *Sci. Adv.*, **7**, eabl4064.
16. Travis,B.A., Ramsey,K.M., Prezioso,S.M., Tallo,T., Wandzilak,J.M., Hsu,A., Borgnia,M., Bartesaghi,A., Dove,S.L., Brennan,R.G. *et al.* (2021) Structural basis for virulence activation of *Francisella tularensis*. *Mol. Cell*, **81**, 139–152.
17. Yang,Y., Liu,C., Zhou,W., Shi,W., Chen,M., Zhang,B., Schatz,D.G., Hu,Y. and Liu,B. (2021) Structural visualization of transcription activated by a multidrug-sensing MerR family regulator. *Nat. Commun.*, **12**, 2702.
18. Fang,C., Philips,S.J., Wu,X., Chen,K., Shi,J., Shen,L., Xu,J., Feng,Y., O'Halloran,T.V. and Zhang,Y. (2021) CueR activates transcription through a DNA distortion mechanism. *Nat. Chem. Biol.*, **17**, 57–64.
19. Fang,C., Li,L., Zhao,Y., Wu,X., Philips,S.J., You,L., Zhong,M., Shi,X., O'Halloran,T.V., Li,Q. *et al.* (2020) The bacterial multidrug resistance regulator BmrR distorts promoter DNA to activate transcription. *Nat. Commun.*, **11**, 6284.
20. Shi,W., Zhang,B., Jiang,Y., Liu,C., Zhou,W., Chen,M., Yang,Y., Hu,Y. and Liu,B. (2021) Structural basis of copper-efflux-regulator-dependent transcription activation. *Iscience*, **24**, 102449.
21. Lilic,M., Darst,S.A. and Campbell,E.A. (2021) Structural basis of transcriptional activation by the *Mycobacterium tuberculosis* intrinsic antibiotic-resistance transcription factor WhiB7. *Mol. Cell*, **81**, 2875–2886.
22. Feng,Y., Zhang,Y. and Ebright,R.H. (2016) Structural basis of transcription activation. *Science*, **352**, 1330–1333.
23. Liu,B., Hong,C., Huang,R.K., Yu,Z. and Steitz,T.A. (2017) Structural basis of bacterial transcription activation. *Science*, **358**, 947–951.
24. Hao,M., Ye,F., Jovanovic,M., Kotta-Loizou,I., Xu,Q., Qin,X., Buck,M., Zhang,X. and Wang,M. (2021) Structures of class I and class II transcription complexes reveal the molecular basis of RamA-dependent transcription activation. *Adv. Sci. (Weinh.)*, **9**, e2103669.
25. Busby,S.J.W. (2019) Transcription activation in bacteria: ancient and modern. *Microbiology (Reading)*, **165**, 386–395.
26. Haakonsen,D.L., Yuan,A.H. and Laub,M.T. (2015) The bacterial cell cycle regulator GcrA is a sigma70 cofactor that drives gene expression from a subset of methylated promoters. *Genes Dev.*, **29**, 2272–2286.
27. Wu,X., Haakonsen,D.L., Sanderlin,A.G., Liu,Y.J., Shen,L., Zhuang,N., Laub,M.T. and Zhang,Y. (2018) Structural insights into the unique mechanism of transcription activation by *Caulobacter crescentus* GcrA. *Nucleic Acids Res.*, **46**, 3245–3256.
28. Adhikari,S., Erill,I. and Curtis,P.D. (2021) Transcriptional rewiring of the GcrA/CcrM bacterial epigenetic regulatory system in closely related bacteria. *PLoS Genet.*, **17**, e1009433.

29. Fioravanti,A., Fumeaux,C., Mohapatra,S.S., Bompard,C., Brilli,M., Frandi,A., Castric,V., Villeret,V., Viollier,P.H. and Biondi,E.G. (2013) DNA binding of the cell cycle transcriptional regulator GcrA depends on N6-adenosine methylation in *Caulobacter crescentus* and other *Alphaproteobacteria*. *PLoS Genet.*, **9**, e1003541.
30. Holtzendorff,J., Hung,D., Brende,P., Reisenauer,A., Viollier,P.H., McAdams,H.H. and Shapiro,L. (2004) Oscillating global regulators control the genetic circuit driving a bacterial cell cycle. *Science*, **304**, 983–987.
31. Fioravanti,A., Fumeaux,C., Mohapatra,S.S., Bompard,C., Brilli,M., Frandi,A., Castric,V., Villeret,V., Viollier,P.H. and Biondi,E.G. (2013) DNA binding of the cell cycle transcriptional regulator GcrA depends on N6-adenosine methylation in *Caulobacter crescentus* and other *Alphaproteobacteria*. *PLoS Genet.*, **9**, e1003541.
32. Mohapatra,S.S., Fioravanti,A., Vandame,P., Spriet,C., Pini,F., Bompard,C., Blossey,R., Valette,O. and Biondi,E.G. (2020) Methylation-dependent transcriptional regulation of crescentin gene (creS) by GcrA in *Caulobacter crescentus*. *Mol. Microbiol.*, **114**, 127–139.
33. Saecker,R.M., Chen,J., Chiu,C.E., Malone,B., Sotiris,J., Ebrahim,M., Yen,L.Y., Eng,E.T. and Darst,S.A. (2021) Structural origins of Escherichia coli RNA polymerase open promoter complex stability. *Proc. Natl. Acad. Sci. U.S.A.*, **118**, e2112877118.
34. Bachhawat,P., Swapna,G.V., Montelione,G.T. and Stock,A.M. (2005) Mechanism of activation for transcription factor PhoB suggested by different modes of dimerization in the inactive and active states. *Structure*, **13**, 1353–1363.
35. Weldon,J.E., Rodgers,M.E., Larkin,C. and Schleif,R.F. (2007) Structure and properties of a truely apo form of AraC dimerization domain. *Proteins*, **66**, 646–654.
36. Maddocks,S.E. and Oyston,P.C.F. (2008) Structure and function of the LysR-type transcriptional regulator (LTTR) family proteins. *Microbiology (Reading)*, **154**, 3609–3623.
37. Friedman,L.J., Mumm,J.P. and Gelles,J. (2013) RNA polymerase approaches its promoter without long-range sliding along DNA. *Proc. Natl. Acad. Sci. U.S.A.*, **110**, 9740–9745.
38. Stracy,M., Lesterlin,C., Garza de Leon,F., Uphoff,S., Zawadzki,P. and Kapanidis,A.N. (2015) Live-cell superresolution microscopy reveals the organization of RNA polymerase in the bacterial nucleoid. *Proc. Natl. Acad. Sci. U.S.A.*, **112**, E4390–E4399.
39. Wang,F., Redding,S., Finkelstein,I.J., Gorman,J., Reichman,D.R. and Greene,E.C. (2013) The promoter-search mechanism of *Escherichia coli* RNA polymerase is dominated by three-dimensional diffusion. *Nat. Struct. Mol. Biol.*, **20**, 174–181.
40. Chen,X., Qi,Y., Wu,Z., Wang,X., Li,J., Zhao,D., Hou,H., Li,Y., Yu,Z., Liu,W. *et al.* (2021) Structural insights into preinitiation complex assembly on core promoters. *Science*, **372**, eaba8490.