nature structural & molecular biology

Article

https://doi.org/10.1038/s41594-023-01149-7

DNA looping mediates cooperative transcription activation

Received: 21 October 2022

Accepted: 4 October 2023

Published online: 04 January 2024

Check for updates

Shu-Jing Han[1], Yong-Liang Jiang[1] ✉, Lin-Lin You[2], Li-Qiang Shen[2], Xiaoxian Wu[2], Feng Yang[1], Ning Cui[1], Wen-Wen Kong[1], Hui Sun[1], Ke Zhou[1], Hui-Chao Meng[1], Zhi-Peng Chen[1], Yuxing Chen[1], Yu Zhang[2] ✉ & Cong-Zhao Zhou[1] ✉

Transcription factors respond to multilevel stimuli and co-occupy promoter regions of target genes to activate RNA polymerase (RNAP) in a cooperative manner. To decipher the molecular mechanism, here we report two cryo-electron microscopy structures of *Anabaena* transcription activation complexes (TACs): NtcA-TAC composed of RNAP holoenzyme, promoter and a global activator NtcA, and NtcA–NtcB-TAC comprising an extra context-specific regulator, NtcB. Structural analysis showed that NtcA binding makes the promoter DNA bend by ~50°, which facilitates RNAP to contact NtcB at the distal upstream NtcB box. The sequential binding of NtcA and NtcB induces looping back of promoter DNA towards RNAP, enabling the assembly of a fully activated TAC bound with two activators. Together with biochemical assays, we propose a 'DNA looping' mechanism of cooperative transcription activation in bacteria.

Multiple transcription factors act in concert via co-occupying the promoter regions and respond to different input signals to precisely control the transcription output of downstream genes in both prokaryotes and eukaryotes[1–5]. In bacteria, one global regulator and one context-specific regulator usually function as a pair to activate gene transcription in a cooperative manner[1,3,4]. Despite structures of TACs containing either class I or class II transcription activators having been reported[6–8], the underlying mechanism of cooperative transcription activation remains unclear due to the lack of structural information for a TAC comprising two activators.

Carbon/nitrogen (C/N) balance control is critical to maintaining cellular homeostasis for all organisms[9]. As an ancient photoautotrophic bacterium, cyanobacteria have evolved a finely regulated signal transduction network to keep pace with nitrogen assimilation with CO_2 fixation[9–12]. The global regulator NtcA in the catabolite activator protein (CAP) family plays a central role in maintaining the C/N balance through sensing the central metabolite 2-oxoglutarate (2-OG), which tightly controls the expression of more than 120 genes that stimulate nitrogen metabolism in *Anabaena* sp. PCC 7120 (*Anabaena* for short)[13–18]. Cyanobacteria mainly utilize the nitrate assimilation pathway encoded by *nirA, nrtABCD and narB* genes (termed *nirA* operon in *Anabaena*) to import and reduce nitrate[10,11,19]. The expression of *nirA* operon is maintained at a basal level by the global regulator NtcA, but significantly boosted by the context-specific activator NtcB (a LysR-type transcriptional regulator; LTTR) when the nitrogen source is switched from ammonium to nitrate[20–24]. Therefore, the regulation of *nirA* operon by NtcA and NtcB serves as an excellent model for studying cooperative transcription activation in bacteria.

Results

NtcA and NtcB cooperatively regulate the expression of *nirA* operon

We first verified that NtcA and NtcB activate transcription of the promoter of the *nirA* operon (P*nirA*) in a cooperative manner. The in vitro transcription activity assays showed that 2-OG-bound NtcA increases the transcription activity of RNAP on P*nirA* (Extended Data Fig. 1a); by contrast, NtcB alone has no effect on P*nirA* transcription (Extended Data Fig. 1b). However, NtcB substantially augments P*nirA* transcription in

[1]School of Life Sciences and Biomedical Sciences and Health Laboratory of Anhui Province, Division of Life Sciences and Medicine, University of Science & Technology of China, Hefei, China. [2]Key Laboratory of Synthetic Biology, CAS Center for Excellence in Molecular Plant Sciences, Shanghai Institute of Plant Physiology and Ecology, Chinese Academy of Sciences, Shanghai, China. ✉e-mail: jyl@ustc.edu.cn; yzhang@cemps.ac.cn; zcz@ustc.edu.cn

https://doi.org/10.1038/s41594-023-01149-7

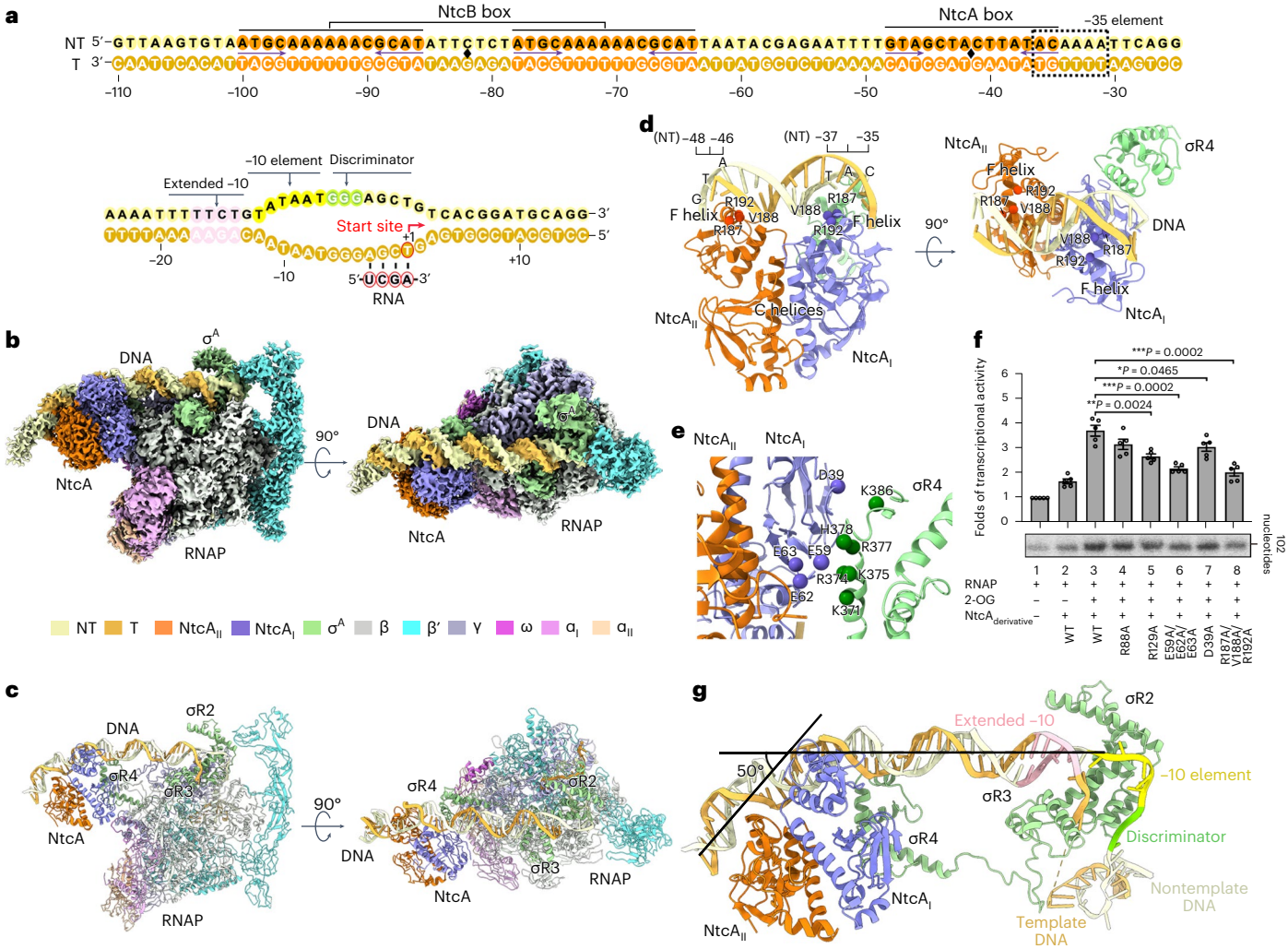


Fig. 1 | The structure of NtcA-TAC. a, The nucleic acid scaffold used for cryo-EM structure determination of NtcA-TAC and NtcA–NtcB-TAC. The NtcA and NtcB binding sites are highlighted and labeled with NtcA box and NtcB box, respectively. The palindromic sequences are underlined by purple arrows. The −35 element is highlighted in a dashed box. **b**,**c**, The cryo-EM map (**b**) and the structure model (**c**) of NtcA-TAC presented in two view orientations. The RNAP subunits, NtcA dimer and the promoter are colored differently, as shown in the color key. T, template; NT, nontemplate. The σR2, σR3 and σR4 represent the region 2, region 3 and region 4 of σ^A, respectively. **d**, NtcA and σR4 recognize partially overlapped NtcA box and the −35 element, respectively. NtcA dimer and σR4 are represented as cartoon. Cα atoms of the key interaction residues are represented as spheres. **e**, Interactions between NtcA and σR4. Cα atoms of the interface residues are shown as spheres. **f**, In vitro transcription activation results using a modified P*nirA* as template DNA and WT/mutant NtcA. Data are presented as mean ± s.e.m., $n = 5$ biologically independent experiments. Two-tailed unpaired *t*-tests. No *P* value adjustments for multiple comparisons. $^*P < 0.05$, $^{**}P < 0.01$ and $^{***}P < 0.001$ indicate significant difference compared to lane 3. The lower panel shows the representative gel image. RNA transcripts, 102 nucleotides. WT, wild type. **g**, The binding of NtcA dimer induces a 50° bending of the promoter DNA at the −35 element.

the presence of 2-OG-bound NtcA (Extended Data Fig. 1c). Our results show that the transcription activation activity of NtcB relies on NtcA, confirming the cooperative model of action by NtcA and NtcB[10,20–23].

Overall architecture of NtcA-TAC

To understand how NtcA activates transcription, we reconstituted the NtcA-TAC using *Anabaena* RNAP, σ^A, 2-OG-bound NtcA and the 125-base pair (bp) P*nirA*-derived DNA comprising the consensus binding motifs of NtcA (−48 to −35 with respect to the transcription start site, centered at −41.5) and NtcB (−100 to −64, centered at −82), also termed NtcA and NtcB boxes (Fig. 1a and Extended Data Fig. 1d,e). The cryo-EM map of NtcA-TAC reconstructed at 3.6 Å (Table 1 and Extended Data Fig. 2) shows unambiguous signals for 66-bp double-stranded (ds) DNA (−53 to +13) with a partially resolved transcription bubble, RNAP and NtcA dimer (Fig. 1b,c and Extended Data Fig. 3a). The RNAP and the core promoter region (−36 to +13) adopt the similar conformation and make essentially the same interactions with each other, as observed in the previously reported open conformation of RNAP–promoter DNA (RPo) structure (Extended Data Fig. 3a)[25].

NtcA functions as a class II activator on the P*nirA* promoter. The proximal protomer ($NtcA_I$) of NtcA dimer makes interactions with region 4 of σ^A (σR4) through one surface patch of NtcA that corresponds to the activating region 3 of *Escherichia coli* CAP (Fig. 1d,e and Extended Data Fig. 3b,c)[26,27]. The interactions involve residues Asp39, Glu59, Glu62 and Glu63 of NtcA and residues Lys371, Arg374, Lys375, Arg377, His378 and Lys386 of σR4 (Fig. 1e and Extended Data Fig. 3c). Mutating these interface residues on NtcA substantially impaired its transcription activation activity, indicating an important role of NtcA–σR4 interaction on the transcription activation activity of NtcA (Fig. 1f). Intriguingly, different from typical class II activators[6,8,27], NtcA makes interaction with neither the N-terminal (αNTD) nor the C-terminal domain (αCTD) of the α subunit of RNAP.

σR4 and NtcA recognize the −35 element and NtcA box, respectively, which are partially overlapped (Fig. 1a,d and Extended Data

Fig. 3b,c). The σR4 loosely contacts the −35 element, and the helix–turn–helix (HTH) module of σR4 does not reach the major groove as deeply as it does in the transcription initiation complex structure in *Synechocystis* sp. PCC 6803 (Extended Data Fig. 3d)[25], probably due to change of the dsDNA trajectory around the −35 element caused by the NtcA–DNA interactions, explaining a degenerated −35 sequence of P*nirA* promoter (Fig. 1a and Extended Data Fig. 1d,e). The degenerate −35 motif and the weakened recognition by σR4 have also been observed in other class II TAC structures[6,28]. The DNA-binding domains (NtcA-DBDs) of the NtcA dimer recognize the palindromic motif of the 14-bp NtcA box ($^{-48}$GTA-N$_8$-TAC^{-35}) centered at −41.5. The 'F helix' and the side wing of the winged HTH module of NtcA-DBD insert into the major and minor grooves of the half palindromic NtcA box, respectively (Fig. 1d). Residues Arg187, Val188 and Arg192 of NtcA-DBD make direct interactions with DNA (Extended Data Fig. 3e). Alanine substitution of these residues not only greatly reduced the binding affinity of NtcA towards the promoter DNA (Extended Data Fig. 3f), but also lowered its transcription activation activity (Fig. 1f). Binding of NtcA dimer to the NtcA box enforces an ~50° bending of DNA towards RNAP (Fig. 1g). This bending changes the trajectory of upstream dsDNA, shortens the spatial distance between RNAP and upstream DNA and eventually makes RNAP accessible to the transcription factor that binds to the distal upstream DNA.

Overall architecture of NtcA–NtcB-TAC

To understand how NtcB cooperates with NtcA to activate transcription, we determined the structure of a TAC comprising both NtcA and NtcB (NtcA–NtcB-TAC). The gel-shift assays showed that the sequential addition of NtcA and NtcB gradually slowed the migration of the RNAP–promoter complex, suggesting the formation of a larger complex (Extended Data Fig. 4). The reconstituted complex of NtcA–NtcB-TAC that comprises *Anabaena* RNAP, σ^A, 2-OG-bound NtcA, NtcB and the 125-bp P*nirA*-derived promoter DNA was subsequently applied to cryo-EM single-particle analysis. The overall 4.5-Å cryo-EM map of NtcA–NtcB-TAC (Table 1 and Extended Data Fig. 5) shows unambiguous signals for RNAP, σ^A, the NtcA dimer, the NtcB tetramer and the ~112-bp promoter DNA with a partially resolved transcription bubble (Fig. 2a,b). The structural model of NtcA–NtcB-TAC was built by sequentially docking NtcA-TAC, the 2.4-Å crystal structure of the dimeric effector-binding domain of NtcB (NtcB-EBD) determined in this study (Table 2 and Extended Data Fig. 6a,b) and a predicted model of the DNA-binding domain of NtcB (NtcB-DBD) into the cryo-EM map. The structure shows that RNAP core promoter interactions in the NtcA–NtcB-TAC remain unchanged compared to the structure of NtcA-TAC, suggesting that binding of NtcB does not alter RPo conformation (Extended Data Fig. 7a). The interactions of NtcA with RNAP and DNA are essentially the same in both NtcA-TAC and NtcA–NtcB-TAC structures, suggesting that NtcB binding does not affect the binding mode of NtcA (Extended Data Fig. 7b).

In the NtcA–NtcB-TAC structure, NtcA and NtcB respectively interact with their boxes and slightly contact each other (Fig. 2a,b). Beyond the interactions between RNAP holoenzyme and the core region of promoter, NtcA binds to σR4 and the NtcA box, whereas NtcB makes interactions with RNAP-αNTD and the NtcB box (Figs. 1 and 2). The electrophoretic mobility shift assays (EMSA) show that both NtcA and NtcB bind to P*nirA* in a sequence-specific manner (Extended Data Fig. 8a,b). Moreover, mutating either NtcA or NtcB box did not affect the DNA-binding affinity of the other transcription factor (Extended Data Fig. 8c,d). Our results suggest that NtcA and NtcB respectively bind to their boxes in an independent manner, different from the mutually dependent DNA-binding mode of many eukaryotic transcription activators[29–31].

The NtcB interactions of the promoter DNA and RNAP

The LTTR family regulator, NtcB, adopts a tetrameric structure that consists of two compact (NtcB$_{II}$ and NtcB$_{III}$) and two extended subunits (NtcB$_I$ and NtcB$_{IV}$), respectively (Fig. 2c). Two NtcB-EBD dimers (NtcB-EBD$_I$/NtcB-EBD$_{III}$ and NtcB-EBD$_{II}$/NtcB-EBD$_{IV}$) are flanked at the two different sides of promoter DNA and separated ~80 Å in distance from each other, while two NtcB-DBD dimers (NtcB-DBD$_{III}$/NtcB-DBD$_{IV}$ and NtcB-DBD$_I$/NtcB-DBD$_{II}$) are lined up along the promoter DNA (Fig. 2c,d). Generally, the NtcB tetramer adopts an expanded conformation that enables binding to both the promoter DNA and RNAP via DBDs and EBDs, respectively.

The NtcA–NtcB-TAC structure reveals how NtcB interacts with DNA (Fig. 2c,d). Similar to typical LTTRs[24,32], each of the two NtcB-DBD

Table 1 | Cryo-EM data collection, refinement and validation statistics

	NtcA-TAC (EMD-34476), (PDB 8H40)	NtcA–NtcB-TAC (EMD-34475), (PDB 8H3V)
Data collection and processing		
Magnification	×81,000	×81,000
Voltage (kV)	300	300
Electron exposure (e$^-$/Å^2)	50	50
Defocus range (μm)	−1.2 to −2.2	−1.2 to −2.2
Pixel size (Å)	1.07	1.07
Symmetry imposed	*C*1	*C*1
Initial particle images (no.)	566,246	851,868
Final particle images (no.)	45,239	65,256
Map resolution (Å)	3.6	4.5
FSC threshold	0.143	0.143
Map resolution range (Å)	3–7.5	3.7–10
Refinement		
Initial model used (PDB code)	8GZG	8GZG
Model resolution (Å)	3.5	4.5
FSC threshold	0.143	0.143
Model resolution range (Å)	3–7.5	3.7–10
Map sharpening *B* factor (Å^2)	−96.62	−246.01
Model composition		
Non-hydrogen atoms	34,919	46,335
Protein residues	4,130	5,326
Ligands	0	0
B factors (Å^2)		
Protein	184.13	34.68
Nucleotide	209.72	130.41
Ligand	0	0
R.m.s. deviations		
Bond lengths (Å)	0.073	0.053
Bond angles (°)	4.360	4.061
Validation		
MolProbity score	4.11	4.29
Clashscore	175.10	179.18
Poor rotamers (%)	4.55	7.60
Ramachandran plot		
Favored (%)	60.4	60.7
Allowed (%)	35.8	33.8
Disallowed (%)	3.8	5.5

https://doi.org/10.1038/s41594-023-01149-7

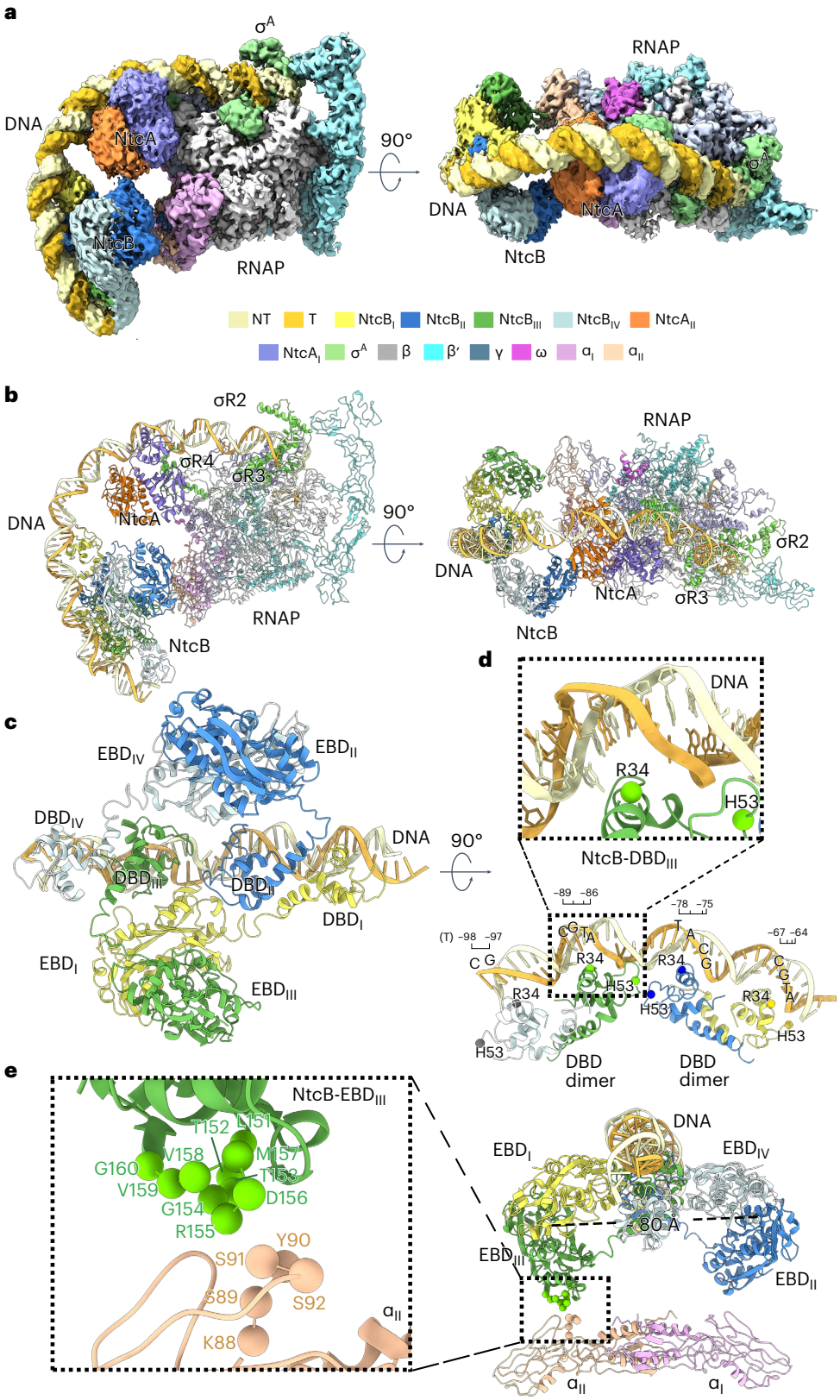


Fig. 2 | The structure of NtcA–NtcB-TAC. a,**b**, The cryo-EM map (**a**) and the structure model (**b**) of NtcA–NtcB-TAC are presented in two view orientations. The RNAP subunits, NtcA and NtcB protomers are colored as shown in the color key. **c**, The NtcB tetramer binds to the NtcB box. Two NtcB-EBD dimers (NtcB-EBD$_I$/NtcB-EBD$_{III}$ and NtcB-EBD$_{II}$/NtcB-EBD$_{IV}$) are flanked at different sides of the promoter DNA. Two NtcB-DBD$_{III}$/NtcB-DBD$_{IV}$ and NtcB-DBD$_I$/NtcB-DBD$_{II}$ dimers are lined up along the promoter DNA. T, template; NT, nontemplate. **d**, The interactions between NtcB-DBDs and NtcB box with a zoomed-in view shown as an inset. The interface residues are shown as spheres. The positions of the conserved recognition motifs are highlighted. **e**, The interactions between NtcB-EBD and RNAP-αNTD with a zoomed-in view shown as an inset. The potential interface residues are shown as spheres.

https://doi.org/10.1038/s41594-023-01149-7

Table 2 | Data collection and refinement statistics

	NtcB-EBD (PDB 8H3Z)
Data collection	
Space group	$P2_12_12_1$
Cell dimensions	
a, *b*, *c* (Å)	59.697, 68.410, 114.310
α, *β*, *γ* (°)	90.00, 90.00, 90.00
Resolution (Å)	44.12-2.40 (2.49-2.40)
R_{merge}	0.124 (0.574)
$I / \sigma I$	20.7 (2.8)
Completeness (%)	99.0 (97.0)
Redundancy	8.3 (5.8)
Refinement	
Resolution (Å)	50.00-2.40
No. reflections	17,649
R_{work} / R_{free}	23.4/28.6
No. atoms	
Protein	3,187
Ligand/ion	25
Water	105
B factors	
Protein	37.86
Ligand/ion	63.59
Water	35.38
R.m.s. deviations	
Bond lengths (Å)	0.013
Bond angles (°)	1.576

dimers is organized in a head-to-head manner to recognize the NtcB box ($^{-100}$ATGC-N$_7$-GCAT-N$_7$-ATGC-N$_7$-GCAT^{-64}) that comprises two palindromic motifs interspaced by 7 bp (Figs. 1a and 2c,d and Extended Data Fig. 1d,e). The NtcB-DBD$_{III}$/NtcB-DBD$_{IV}$ dimer recognizes the distal palindromic box, while the NtcB-DBD$_I$/NtcB-DBD$_{II}$ dimer interacts with the proximal box (Fig. 2c,d). Further structural analysis showed that residues Arg34 and His53 of NtcB-DBD, which are conserved in LTTRs[24,33], directly interact with DNA (Fig. 2d and Extended Data Fig. 9a). Alanine substitution of these residues greatly reduced the binding affinity of NtcB towards the promoter (Extended Data Fig. 9b) and decreased its transcription activation activity (Extended Data Fig. 9c), confirming the essential roles of these residues.

The NtcA–NtcB-TAC structure also reveals how NtcB interacts with RNAP. Compared to the previously reported structures of LTTRs[24,34], the two EBD dimers of NtcB are distantly separated, whereas the two DBD dimers closely approach each other (Fig. 2d,e). The two separated EBD dimers make interactions with the two RNAP α subunits via NtcB-EBD$_{II}$ and NtcB-EBD$_{III}$, respectively (Fig. 2e and Extended Data Fig. 9d). In the NtcB-EBD$_I$/NtcB-EBD$_{III}$ dimer, a small surface patch (residues Leu151–Gly160) of EBD$_{III}$ approaches the N-terminal domain of the RNAP α_{II} subunit (RNAP-αNTD$_{II}$) and makes interactions most likely with a loop region (residues Lys88–Ser92) of RNAP-αNTD$_{II}$ (Fig. 2e and Extended Data Fig. 9d). Similar interaction networks are also observed for the NtcB-EBD$_{II}$/NtcB-EBD$_{IV}$ dimer and RNAP-αNTD$_I$ (Fig. 2e and Extended Data Fig. 9d). The interactions between RNAP-αNTD and NtcB-EBD are critical, as deletion of the interacting patch (residues Leu151–Gly160) of NtcB substantially impaired the cooperative transcription activation activity (Extended Data Fig. 9c).

The 'DNA looping' model

The binding of NtcA and NtcB to the promoter enforces a DNA bending of ~50° and ~70°, respectively, at their binding boxes (Fig. 3a). Moreover, NtcB binding to RNAP triggers an extra bending of ~20° between NtcA and NtcB boxes. Therefore, the sequential binding of NtcA and NtcB induces a total DNA bending of ~140°, resulting in the looping back of the upstream promoter DNA in the NtcA–NtcB-TAC structure (Fig. 3a). This bending of 140° enables the concomitant binding of NtcA and NtcB with both RNAP and promoter DNA. The 15-bp spacer between the binding boxes of NtcA and NtcB makes it possible for the NtcB tetramer and NtcA dimer to bind to the promoter DNA on the same side that faces RNAP (Figs. 1a and 3a and Extended Data Fig. 1d,e). Indeed, either extending or shortening the spacer by 5 bp completely abolished the cooperative activation activity of NtcB (Extended Data Fig. 10).

These structural and biochemical analyses enabled us to propose a 'DNA looping' mechanism underlying the cooperative activation by NtcA and NtcB, in which the NtcB–RNAP interaction is enabled by the proximal promoter bending induced by the NtcA–RNAP interaction (Fig. 3b,c). The NtcB box is located at the distal upstream of the core promoter region, and thereby the tripartite interaction among NtcB, the promoter DNA, and RNAP-αNTD is energetically costly, as it requires a DNA curving of ~140°. Our structures show that NtcA bends the upstream DNA around its binding box and changes the trajectory of the upstream promoter towards RNAP (Fig. 3b). The NtcA-induced DNA bending significantly decreases the spatial distance between NtcB and RNAP-αNTD, making their interaction become possible. Moreover, it lowers the energetic barrier of DNA curving required for NtcB–RNAP interaction (Fig. 3b,c). The extensive interactions among the activators, RNAP and promoter DNA in the NtcA–NtcB-TAC structure are expected to stabilize the RNAP–DNA complex, as observed in other class I and II bacterial transcription activation complexes[6,7,35,36].

To test this hypothesis, we performed molecular beacon assays to investigate whether NtcA/NtcB affects the interactions between RNAP and a P*nirA* derivative (−105 to −11) containing the binding motifs of NtcA and NtcB, the −35 element, and the −11 A overhang of the non-template DNA. We omitted the downstream DNA region to prevent interference from dsDNA unwinding. The results (Fig. 3d) showed that NtcA alone increases the RNAP–DNA affinity by approximately twofold (14 nM versus 6 nM), while the presence of NtcB and NtcA together increases the RNAP–DNA affinity by approximately sevenfold (14 nM versus 2 nM). NtcB alone had no effect. These findings support the cooperative activation model of NtcA and NtcB and suggest that NtcA and NtcB activate transcription by promoting RNAP–promoter DNA closed complex (RPc) formation. However, we cannot rule out the possibility that the two activators might also facilitate RPo formation. The increased interaction between RNAP and dsDNA might help RNAP open the SI3-σ arch, enabling subsequent loading of the promoter DNA[25]. In summary, our results indicate that NtcA and NtcB enhance the affinity between RNAP and dsDNA, likely promoting RPc formation.

Discussion

Our 'DNA looping' model suggests that a transcription factor that bends the proximal upstream promoter DNA facilitates the interaction between RNAP and the second transcription factor bound at the distal upstream regions. The presence of binding sites of both class II and class I activators has been observed on many bacterial σ^{70}-dependent promoters, suggesting the prevalence of such a mode of regulation in σ^{70}-dependent transcription[1–5,37,38]. Moreover, the 'DNA looping' mechanism has long been hypothesized in σ^{54}-dependent transcription, in which the action of enhancer-binding protein requires DNA looping by bacterial histone-like proteins[1,39]. This model has also been hypothesized in eukaryotic RNA polymerase II transcription, in which the enhancer and promoter are brought together by DNA

https://doi.org/10.1038/s41594-023-01149-7

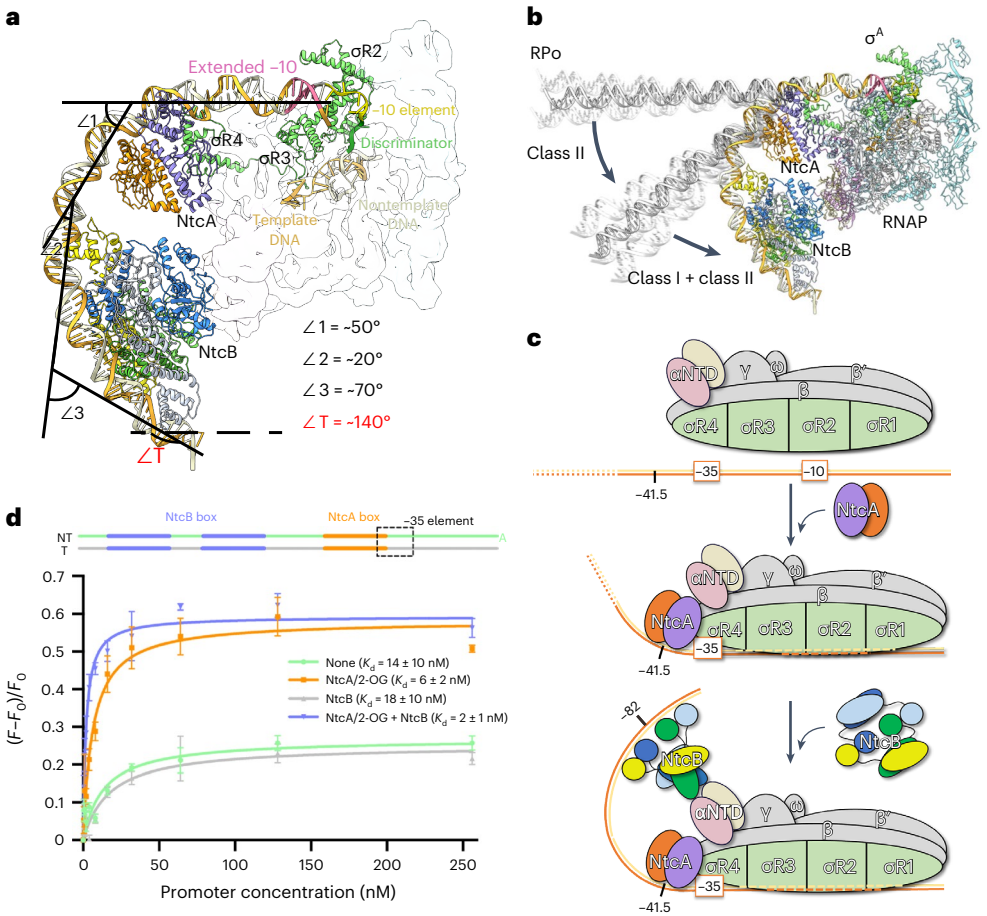


Fig. 3 | The 'DNA looping' model of cooperative transcription activation by NtcA and NtcB. a, The upstream DNA is curved at positions −41.5 (~50°), −61 (~20°), −82 (~70°), resulting in a total bending (∠T) of ~140°. The promoter DNA, NtcA and NtcB protomers are colored as in Fig. 2. **b**, Comparison of the upstream DNA trajectory of RPo, NtcA-TAC and NtcA–NtcB-TAC. The bending of the promoter DNA is sequentially induced by the class II activator NtcA and the class I activator NtcB. The modeled promoter DNAs in flexible motion are colored in white. **c**, The proposed 'DNA looping' model of cooperative transcription activation by NtcA and NtcB. The interactions among NtcA, RNAP and promoter DNA induce a DNA bending by ~50°, which enables RNAP to interact with NtcB located to the distal upstream promoter region. The sequential binding of NtcA and NtcB induces looping back of the promoter DNA towards RNAP, eventually enabling the assembly of a fully activated TAC bound with two activators. **d**, The molecular beacon assays show that the sequential binding of NtcA and NtcB increases the affinity of promoter dsDNA to *Anabaena* RNAP holoenzyme. The DNA schematic used in the beacon assays is shown at the top. The K_d value was obtained from three independent repeats. The data are presented as mean ± s.e.m. T, template; NT, nontemplate.

looping forming proteins, topologically associating domains, as well as condensates[2,29,40–42]. Our findings suggest a general means for transcription cooperative activation shared in bacteria and eukaryotes. Moreover, our work provides a detailed structural example of 'DNA looping'-mediated cooperative transcription activation.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41594-023-01149-7.

References

1. Browning, D. F. & Busby, S. J. Local and global regulation of transcription initiation in bacteria. *Nat. Rev. Microbiol.* **14**, 638–650 (2016).
2. Cramer, P. Organization and regulation of gene transcription. *Nature* **573**, 45–54 (2019).
3. Browning, D. F. & Busby, S. J. The regulation of bacterial transcription initiation. *Nat. Rev. Microbiol.* **2**, 57–65 (2004).
4. Lee, D. J., Minchin, S. D. & Busby, S. J. Activating transcription in bacteria. *Annu. Rev. Microbiol.* **66**, 125–152 (2012).
5. Busby, S. J. W. Transcription activation in bacteria: ancient and modern. *Microbiology* **165**, 386–395 (2019).
6. Feng, Y., Zhang, Y. & Ebright, R. H. Structural basis of transcription activation. *Science* **352**, 1330–1333 (2016).
7. Liu, B., Hong, C., Huang, R. K., Yu, Z. & Steitz, T. A. Structural basis of bacterial transcription activation. *Science* **358**, 947–951 (2017).
8. Shi, W., Jiang, Y., Deng, Y., Dong, Z. & Liu, B. Visualization of two architectures in class-II CAP-dependent transcription activation. *PLoS Biol.* **18**, e3000706 (2020).
9. Forchhammer, K. & Selim, K. A. Carbon/nitrogen homeostasis control in cyanobacteria. *FEMS Microbiol. Rev.* **44**, 33–53 (2020).
10. Ohashi, Y. et al. Regulation of nitrate assimilation in cyanobacteria. *J. Exp. Bot.* **62**, 1411–1424 (2011).
11. Flores, E., Frias, J. E., Rubio, L. M. & Herrero, A. Photosynthetic nitrate assimilation in cyanobacteria. *Photosynth. Res.* **83**, 117–133 (2005).
12. Flores, E. & Herrero, A. Nitrogen assimilation and nitrogen control in cyanobacteria. *Biochem. Soc. Trans.* **33**, 164–167 (2005).

13. Zhao, M. X. et al. Structural basis for the allosteric control of the global transcription factor NtcA by the nitrogen starvation signal 2-oxoglutarate. *Proc. Natl Acad. Sci. USA* **107**, 12487–12492 (2010).
14. Llacer, J. L. et al. Structural basis for the regulation of NtcA-dependent transcription by proteins PipX and PII. *Proc. Natl Acad. Sci. USA* **107**, 15397–15402 (2010).
15. Picossi, S., Flores, E. & Herrero, A. ChIP analysis unravels an exceptionally wide distribution of DNA binding sites for the NtcA transcription factor in a heterocyst-forming cyanobacterium. *BMC Genomics* **15**, 22 (2014).
16. Mitschke, J. et al. Dynamics of transcriptional start site selection during nitrogen stress-induced cell differentiation in *Anabaena* sp. PCC7120. *Proc. Natl Acad. Sci. USA* **108**, 20130–20135 (2011).
17. Dominguez-Martin, M. A. et al. Differential NtcA responsiveness to 2-oxoglutarate underlies the diversity of C/N balance regulation in *Prochlorococcus*. *Front. Microbiol.* **8**, 2641 (2017).
18. Giner-Lamia, J. et al. Identification of the direct regulon of NtcA during early acclimation to nitrogen starvation in the cyanobacterium *Synechocystis* sp. PCC 6803. *Nucleic Acids Res.* **45**, 11800–11820 (2017).
19. Imamura, S. et al. Nitrate assimilatory genes and their transcriptional regulation in a unicellular red alga *Cyanidioschyzon merolae*: genetic evidence for nitrite reduction by a sulfite reductase-like enzyme. *Plant Cell Physiol.* **51**, 707–717 (2010).
20. Lopatovskaia, K. V., Seliverstov, A. V. & Liubetskii, V. A. NtcA- and NtcB-regulons in cyanobacteria and Rhodophyta chloroplasts. *Mol. Biol.* **45**, 570–574 (2011).
21. Frias, J. E., Flores, E. & Herrero, A. Activation of the *Anabaena nir* operon promoter requires both NtcA (CAP family) and NtcB (LysR family) transcription factors. *Mol. Microbiol.* **38**, 613–625 (2000).
22. Aichi, M. & Omata, T. Involvement of NtcB, a LysR family transcription factor, in nitrite activation of the nitrate assimilation operon in the cyanobacterium *Synechococcus* sp. strain PCC 7942. *J. Bacteriol.* **179**, 4671–4675 (1997).
23. Aichi, M., Takatani, N. & Omata, T. Role of NtcB in activation of nitrate assimilation genes in the cyanobacterium *Synechocystis* sp. strain PCC 6803. *J. Bacteriol.* **183**, 5840–5847 (2001).
24. Maddocks, S. E. & Oyston, P. C. F. Structure and function of the LysR-type transcriptional regulator (LTTR) family proteins. *Microbiology* **154**, 3609–3623 (2008).
25. Shen, L. Q. et al. An SI3-σ arch stabilizes cyanobacteria transcription initiation complex. *Proc. Natl Acad. Sci. USA* **120**, e2219290120 (2023).
26. Lawson, C. L. et al. Catabolite activator protein: DNA binding and transcription activation. *Curr. Opin. Struct. Biol.* **14**, 10–20 (2004).
27. Busby, S. & Ebright, R. H. Transcription activation by catabolite activator protein (CAP). *J. Mol. Biol.* **293**, 199–213 (1999).
28. Blanco, A. G., Canals, A., Bernues, J., Sola, M. & Coll, M. The structure of a transcription activation subcomplex reveals how σ^{70} is recruited to PhoB promoters. *EMBO J.* **30**, 3776–3785 (2011).
29. Panigrahi, A. & O'Malley, B. W. Mechanisms of enhancer action: the known and the unknown. *Genome Biol.* **22**, 108 (2021).
30. Haberle, V. & Stark, A. Eukaryotic core promoters and the functional basis of transcription initiation. *Nat. Rev. Mol. Cell Biol.* **19**, 621–637 (2018).
31. Sikorski, T. W. & Buratowski, S. The basal initiation machinery: beyond the general transcription factors. *Curr. Opin. Cell Biol.* **21**, 344–351 (2009).
32. Giannopoulou, E. A. et al. Crystal structure of the full-length LysR-type transcription regulator CbnR in complex with promoter DNA. *FEBS J.* **288**, 4560–4575 (2021).
33. Alanazi, A. M., Neidle, E. L. & Momany, C. The DNA-binding domain of BenM reveals the structural basis for the recognition of a T-N_{11}-A sequence motif by LysR-type transcriptional regulators. *Acta Crystallogr. D Biol. Crystallogr.* **69**, 1995–2007 (2013).
34. Jiang, Y. L. et al. Coordinating carbon and nitrogen metabolic signaling through the cyanobacterial global repressor NdhR. *Proc. Natl Acad. Sci. USA* **115**, 403–408 (2018).
35. Mazumder, A. & Kapanidis, A. N. Recent advances in understanding σ70-dependent transcription initiation mechanisms. *J. Mol. Biol.* **431**, 3947–3959 (2019).
36. Hao, M. et al. Structures of class I and class II transcription complexes reveal the molecular basis of RamA-dependent transcription activation. *Adv. Sci.* **9**, e2103669 (2022).
37. Holst, B. et al. The cAMP-CRP/CytR nucleoprotein complex in *Escherichia coli*: two pairs of closely linked binding sites for the cAMP-CRP activator complex are involved in combinatorial regulation of the *cdd* promoter. *EMBO J.* **11**, 3635–3643 (1992).
38. Rossiter, A. E. et al. Expression of different bacterial cytotoxins is controlled by two global transcription factors, CRP and Fis, that co-operate in a shared-recruitment mechanism. *Biochem. J.* **466**, 323–335 (2015).
39. Danson, A. E., Jovanovic, M., Buck, M. & Zhang, X. Mechanisms of σ^{54}-dependent transcription initiation and regulation. *J. Mol. Biol.* **431**, 3960–3974 (2019).
40. Weintraub, A. S. et al. YY1 is a structural regulator of enhancer-promoter loops. *Cell* **171**, 1573–1588.e28 (2017).
41. Chen, X. et al. Structures of the human Mediator and Mediator-bound preinitiation complex. *Science* **372**, eabg0635 (2021).
42. Schoenfelder, S. & Fraser, P. Long-range enhancer-promoter contacts in gene expression control. *Nat. Rev. Genet.* **20**, 437–455 (2019).

Methods

Plasmid construction

The plasmid information is summarized in Supplementary Table 1. The genes encoding the α subunit (*rpoA*), β subunit (*rpoB*), γ subunit (*rpoC1*), β′ subunit (*rpoC2*), ω subunit (*rpoZ*), σ factor (*sigA*, σ^A), NtcA (*ntcA*) and NtcB (*ntcB*) were amplified by PCR from *Anabaena* genomic DNA, and their mutants (σ^A_{E201C}, $NtcA_{E59A/E62A/E63A}$, $NtcA_{D39A}$, $NtcA_{R187A/V188A/R192A}$, $NtcB_{\Delta Leu151\text{-}Gly160}$, $NtcB_{R34A/H53A}$, NtcB-EBD) were amplified by PCR from their wild-type plasmids, respectively. *Anabaena rpoA* and *rpoZ* were cloned into pCDFDuet™-1 (Novagen); *rpoB*, *rpoC1* and *rpoC2* were cloned into pETDuet (Novagen) using homologous recombination methods. *Anabaena ntcA*, *ntcB*, σ^A and their derivatives were respectively cloned into the pET28a vector with an N-terminal His-tag.

pEASY/P*nirA* was constructed by ligating pEASY-blunt vector (Transgen Biotech) and a DNA fragment (−148 to +102) comprising the upstream region (−148 to −31; amplified from *Anabaena* genomic DNA) of *nirA* operon promoter and the core promoter region of T5 N25 (−30 to +102; amplified from plasmid pARTaq-*N25-100-TR2)*[43]. The pEASY/P$nirA^{AB+5}$ and pEASY/P$nirA^{AB-5}$ were prepared using pEASY/P*nirA* as template through site-directed mutation.

Protein expression and purification

All proteins except for NtcB were expressed in *E. coli* strain BL21 (DE3) (Novagen). NtcB and its mutants were overexpressed in *E. coli* BL21(DE3) that was cotransformed with the constructed plasmid plus pKY206 containing *groEL/ES* genes encoding chaperonins. Protein expression was induced for 16 h by addition of 0.4 mM IPTG at 18 °C when the optical density at 600 nm reaches 0.6 to 0.8. Cells were collected by centrifugation (11,000*g*, 3 min, 4 °C). Cell pellets of NtcA, NtcB and their mutants were resuspended using buffer A (20 mM Tris–HCl pH 7.9, 1 M NaCl, 5% (v/v) octyl β-D-glucopyranoside, 5 mM β-mercaptoethanol and 1 mM phenylmethyl sulfonyl fluoride) and lysed by 30 min of sonication. After centrifugation at 27,000*g* for 30 min, supernatants containing target proteins were loaded onto a Ni-NTA column (Cytiva), respectively, equilibrated with buffer A. Target proteins were eluted using buffer A with 0.5 M imidazole and further loaded onto a Superdex 200 column (GE Healthcare) pre-equilibrated with buffer A. Target protein fractions were eluted with buffer A, pooled and concentrated to 10 mg ml^{-1}. The cell pellet of NtcB-EBD was resuspended using buffer B (8.1 mM Na_2HPO_4, 1.76 mM KH_2PO_4, 2.67 mM KCl, 136.89 mM NaCl and 5 mM β-mercaptoethanol) and lysed by 30 min of sonication and purified by following the above procedure.

Anabaena RNAP core enzyme was overexpressed in *E. coli* BL21 (DE3) cells carrying pETDuet/*rpoB-rpoC1-rpoC2* and pCDFDuet-1/*rpoA-rpoZ* and purified as described in our previous report[25]. *Anabaena* σ^A and σ^A_{E201C} mutant were overexpressed in *E. coli* BL21 (DE3) cells carrying pET28a/σ^A and pET28a/σ^A_{E201C}, and purified by a similar procedure to that for NtcA and NtcB.

Crystallization and structure determination of NtcB-EBD

Crystals of *Anabaena* NtcB-EBD were grown at 16 °C by sitting-drop vapor diffusion in a 2-µl drop (1 µl of 15 mg ml^{-1} protein solution and 1 µl of reservoir solution comprising 0.1 M MES pH 6.2, 1 M Li_2SO_4). The iodine-substituted crystals were prepared by soaking the crystals with 1 M KI. X-ray diffraction data were collected at 100 K at beamline 19U with a DECTRIS PILATUS 6M detector at the Shanghai Synchrotron Radiation Facility. The diffraction data were integrated and scaled using XDS[44]. The crystal structure of NtcB-EBD was determined by experimental phasing using PHENIX AutoSol[45]. The initial model was built by AutoBuild, refined in REFMAC5 (ref. 46) and manually rebuilt interactively using the program Coot[47]. The final model showed good geometry and was evaluated using MolProbity (http://molprobity.biochem.duke.edu). A list of the parameters for data collection, processing, structure determination and refinement is provided in Table 2.

Nucleic acid scaffold

Nucleic acid scaffold for *Anabaena* NtcA–NtcB-TAC and NtcA-TAC was prepared as follows: nontemplate strand DNA (NT-DNA, 0.73 mM final; Sangon Biotech), template strand DNA (T-DNA, 0.8 mM final; Sangon Biotech) and oligoribonucleotide (5′-UCGA-3′; 7.1 mM; GenScript) in 42.4 µl 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 25 °C in 1 °C steps with 30 s per step using a thermal cycler. The sequences of template and nontemplate DNAs are shown in Fig. 1a and Extended Data Fig. 1e, and listed in Supplementary Table 1.

Electrophoretic mobility shift assays

Cy5-labeled DNA fragments used in EMSA assays for NtcB and NtcA binding are listed in Supplementary Table 1. The reaction mixture (10 µl) includes DNA fragment (0.2 µM, Sangon Biotech), WT/mutant NtcA (0, 0.5, 1, 2, 4 µM) or WT/mutant NtcB (0, 2, 4, 8, 16, 32 µM) in 20 mM Tris–HCl pH 8.0, 100 mM NaCl. The mixture was incubated for 20 min on ice. The solutions were separated by 5% Tris–boric acid (TBE) gel in TBE buffer (89 mM Tris–boric pH 8.3, 2 mM EDTA) and analyzed by Tanon-2500 (Tanon Science & Technology).

For gel-shift assays, the reaction mixture of 7 µl contained 10.5 µM *Anabaena* RNAP–σ^A holoenzyme, 7 µM P*nirA*, 0 or 28 µM NtcA, 0 or 56 µM NtcB in 10 mM HEPES pH 7.5, 100 mM KCl, 5 mM $MgCl_2$, 5 mM DTT. The reactions were incubated for 20 min at room temperature and 1 h in ice. The solutions were separated by 5% Tris–glycine gel in the Tris–glycine buffer (25 mM Tris–HCl, 192 mM glycine pH 8.3), stained using SYBR Gold and analyzed by Tanon-2500 (Tanon Science & Technology).

In vitro transcription assays

To measure the transcription activation activity of *Anabaena* NtcA, reaction mixtures of 20 µl containing 25 nM RNAP–σ^A holoenzyme, 40 mM Tris–HCl pH 8.0, 5 mM $MgCl_2$, 75 mM NaCl, 12.5% (v/v) glycerol and 2.5 mM DTT and 25 nM P*nirA* (prepared by PCR using pEASY/P*nirA* as a template) were incubated for 10 min at 37 °C. NtcA (0, 25, 50, 100, 200, 400 or 800 nM) preincubated with 2-OG at a concentration of 100 times the NtcA concentration was added into the mixture followed by 10-min incubation at 37 °C. The reactions were initiated by adding NTP mixture (100 µM [α-^{32}P]UTP of 0.04 Bq $fmol^{-1}$, 100 µM of ATP, GTP and CTP) and terminated after 10 min at 37 °C by addition of 5 µl of loading buffer (8 M urea, 20 mM EDTA, 0.025% xylene cyanol and 0.025% bromophenol blue). The samples were boiled for 2 min, and immediately cooled on ice for 5 min.

To measure the transcription activation activity by NtcB, reaction mixtures of 20 µl containing 50 nM RNAP–σ^A holoenzyme, 40 mM Tris–HCl pH 8.0, 5 mM $MgCl_2$, 75 mM NaCl, 12.5% (v/v) glycerol, 2.5 mM DTT and 50 nM P*nirA* or its derivatives (prepared by PCR using pEASY/P*nirA*, pEASY/P$nirA^{AB+5}$ and pEASY/P$nirA^{AB-5}$ as templates, respectively) were incubated for 10 min at 37 °C. Then, 0, 50, 100, 200, 400, 800 or 1,600 nM NtcB and 100 nM NtcA preincubated with 1 µM 2-OG were added into the reaction mixture followed by 10-min incubation at 37 °C. The reactions were initiated by adding 100 µM NTP mixture and terminated after 10 min at 37 °C by addition of 5 µl of loading buffer. The samples were boiled for 2 min, and immediately cooled on ice for 5 min.

To measure the transcription activation activity by NtcA mutants, reaction mixtures of 20 µl containing 25 nM RNAP–σ^A holoenzyme, 40 mM Tris–HCl pH 8.0, 5 mM $MgCl_2$, 75 mM NaCl, 12.5% (v/v) glycerol, 2.5 mM DTT and 25 nM P*nirA* (prepared by PCR using pEASY/P*nirA* as template) were incubated for 10 min at 37 °C. Then, 100 nM WT/mutant NtcA preincubated with 1 µM 2-OG was added into the reaction mixture followed by 10-min incubation at 37 °C. To measure the transcription activation activity by NtcB mutants, 50 nM NtcA preincubated with 500 nM 2-OG and 400 nM WT/mutant NtcB were added into the RNAP–σ^A holoenzyme–P*nirA* mixture followed by 10-min incubation at 37 °C. The reactions were initiated by adding 100 µM NTP mixture and

terminated after 10 min at 37 °C by addition of 5 µl of loading buffer. The samples were boiled for 2 min, and immediately cooled in ice for 5 min.

RNA transcripts were separated in 15% (19:1, acrylamide: bisacrylamide) urea–polyacrylamide slab gels in 90 mM Tris–borate (pH 8.0) and 0.2 mM EDTA. The radiograph was obtained by storage-phosphor scanning (Typhoon; GE Healthcare) and the intensity was quantified by the software FUJIFILM Multi Gauge. The DNA templates used for in vitro transcription assays are shown in Extended Data Fig. 1e and listed in Supplementary Table 1.

Protein labeling

Anabaena σ^{A}_{E201C} mutant was reduced by 2 mM DTT and further purified in PBS buffer using a Superdex 200 column (GE Healthcare). The fractions were cooled on ice and quickly incubated with a 5-molar ratio of 5-tetramethylrhodamine and kept in the dark for 2 h. Then, 5 mM DTT was added to terminate the reaction. The unreacted dye was removed using a PD-10 desalting column (GE Healthcare). The labeled protein was pooled and concentrated to 5 mg ml^{-1}. The labeled RNAP holoenzyme used in the molecular beacon assays was reconstituted by mixing *Anabaena* RNAP core enzyme and 5-tetramethylrhodamine-labeled σ^{A}_{E201C} in 10 mM Tris–HCl pH 7.7, 100 mM NaCl, 2 mM DTT and 1% glycerol, in an approximately 1:4 molar ratio.

Molecular beacon assays

The interaction between RNAP and promoter DNA (Extended Data Fig. 1e and Supplementary Table 1) was measured by molecular beacon assays[48–50]. The 100 µl reaction mixtures consisted of *Anabaena* 1 nM RNAP holoenzyme containing σ^{A}_{E201C} labeled with 5-tetramethylrhodamine, the transcription activators as indicated (NtcA, 500 nM; NtcB, 2,000 nM) and 0, 0.5, 1, 2, 4, 8, 16, 32, 64 or 128 nM dsDNA with the −11 A overhang at the nontemplate strand DNA (sequences are listed in Supplementary Table 1), 10 mM Tris–HCl pH 8.0, 200 mM NaCl, 1 mM DTT, 0.025% Tween-20 and 1% glycerol. NtcA was preincubated with 0.5 mM 2-OG. Reaction mixtures were incubated for 15 min at room temperature. Fluorescence emission intensities were recorded using a microplate reader (TECAN; excitation wavelength, 550/10 nM; emission wavelength, 580/20 nM). The equilibrium dissociation constants (K_d) were estimated by nonlinear regression using the equation:

$$F/F_0 = \{B[S]/(K_d + [S])\} + 1 \quad (1)$$

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

Cryo-EM sample preparation and data collection

The NtcA-TAC complex was reconstituted by mixing *Anabaena* RNAP–σ^{A} holoenzyme, NtcA and nucleic acid scaffold in a molar ratio of 1:4:1.3, while NtcA–NtcB-TAC was reconstituted by mixing *Anabaena* RNAP–σ^{A} holoenzyme, NtcA, NtcB and nucleic acid scaffold in a molar ratio of 1:4:10:1.3. Then 5 mM 2-OG was added to the reconstituted mixture and incubated for 2 h on ice.

The NtcA-TAC and NtcA–NtcB-TAC complexes were further purified and cross-linked by gradient fixation (GraFix)[51]. In brief, the solution was added to the top of a continuous density gradient (0–30% (v/v) glycerol and 0.05–0.02% paraformaldehyde) in buffer C (10 mM HEPES pH 7.5, 100 mM KCl, 5 mM $MgCl_2$ and 2 mM DTT). Then, the mixture was centrifuged at 192,230*g* for 16 h at 4 °C using a SW40Ti rotor (Beckman Coulter). The fractions containing target complex were collected and loaded on a Superose 6 10/300 GL column (GE Healthcare) in buffer C to remove glycerol. The fractions of NtcA-TAC and NtcA–NtcB-TAC were pooled and concentrated to ~15 and ~17 mg ml^{-1}, respectively.

The freshly prepared complexes were incubated with 3-([3-cholamidopropyl] dimethylammonio)-2-hydroxy-1-propanesulfonate (CHAPSO, 8 mM; Hampton Research) before grid preparation[52]. Then, a drop of 3 µl was placed on a glow-discharged holey gold grid (UltrAuFoil, R1.2/1.3 Au 400 mesh, Protochips), blotted with Vitrobot Mark IV (FEI) and plunge-frozen into liquid ethane with 100% chamber humidity at 22 °C. The grids were blotted for 0.5 s with a blot force of −1 and a wait time of 5 s.

The NtcA-TAC and NtcA–NtcB-TAC cryo-EM datasets were collected using EPU on a 300 keV Titan Krios (FEI) equipped at the University of Science and Technology of China (USTC). For NtcA-TAC, a total of 2,090 videos (32 frames, each 0.16 s, total dose ~50 $e^{-}/Å^{-2}$) were recorded using a Gatan K3 Summit direct electron detector in super-resolution mode at a nominal magnification of ×81,000 at a pixel size of 0.535 Å and with a defocus range from −1.2 to −2.2 µm. For NtcA–NtcB-TAC, a total of 5,047 videos (40 frames, each 0.16 s, total dose ~50 $e^{-}/Å^{-2}$) were recorded using a Gatan K3 Summit direct electron detector in super-resolution mode at a nominal magnification of ×81,000 at a pixel size of 0.535 Å and with a defocus range from −1.2 to −2.2 µm.

Cryo-EM data processing

All of the videos were motion corrected and dose weighted using MotionCor2 (ref. 53) and were binned twofold to yield a pixel size of 1.07 Å. The defocus values were estimated using CTFFIND4 (ref. 54).

For the dataset of NtcA-TAC, a total of 1,462 NtcA-TAC micrographs were processed for particle autopicking using RELION after removing bad images by manual checking[55]. Approximately 566,246 particles were boxed for two-dimensional classification. All good classes containing 555,921 particles were selected for the first round of three-dimensional (3D) classification with four classes by *C*1 symmetry. A total of 298,068 particles represented the complex containing RNAP and NtcA, whereas the rest of the three classes represented irrelevant junk particles. Approximately ~82.4% of particles were selected for the second-round 3D classification and refinement, yielding a map of NtcA-TAC at 3.4 Å resolution after a PostProcess step as determined by gold standard Fourier shell correlation (FSC) using the 0.143 threshold. However, the cryo-EM map showed weak signal for the upstream portion of NtcA. To obtain a better map near the NtcA region, we performed the third-round 3D classification without alignment and yielded 211,399 particles for the next masked 3D classification focusing on NtcA. Then, ~21.4% particles were re-extracted to perform 3D refinement with solvent flattened FSCs, CTF refinement and Bayesian polishing, yielding a final map at a resolution of 3.6 Å.

For the dataset of NtcA–NtcB-TAC, a total of 402,817 particles were selected after removing bad images and irrelevant junk particles for the first-round 3D classification with four classes by *C*1 symmetry from 5,047 images. Approximately ~16.2% of particles represented the complex containing RNAP, NtcA and NtcB. The particles were imported into CryoSPARC for the next 3D refinement and local refinement, yielding a final map at a resolution of 4.5 Å. To obtain a better map at the region of the NtcA and NtcB binding sites, 253,244 particles were selected to perform 3D classification focusing on the NtcA and NtcB binding regions without alignment. In total, 55,966 particles were selected for the final 3D refinement by imposing *C*1 symmetry, yielding a map at a resolution of 7.6 Å.

Model building and refinement

Model building of NtcA-TAC was performed using Chimera[56] by manually fitting the *Synechocystis* sp. PCC 6803 transcription initiation complex (PDB 8GZG) into the cryo-EM map. NtcA dimer was unambiguously docked into the map using the previously solved crystal structure of *Anabaena* sp. PCC 7120 NtcA (PDB 3LA2)[13]. Model building of the intact NtcA–NtcB-TAC was performed using Chimera by manually fitting the NtcA-TAC into the map. The structural model of full-length NtcB was constructed by a predicted model of NtcB-DBD and the crystal structure of NtcB-EBD (PDB 8H3Z). The chimeric NtcB model was subsequently docked into the map in Chimera. The two

structure models were adjusted in Coot[47], followed by the iterative positional and *B*-factor refinement in real space using PHENIX[45]. The geometry of the final structures was evaluated using MolProbity (http://molprobity.biochem.duke.edu). A list of the parameters of cryo-EM data collection, processing, structure determination and refinement are provided in Table 1.

Statistics and reproducibility

No statistical method was used to predetermine sample size. No data were excluded from the analyses. The experiments were not randomized. The investigators were not blinded to allocation during experiments and outcome assessment.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The crystal structure of NtcB-EBD has been deposited at Protein Data Bank (PDB) under accession code 8H3Z. The cryo-EM maps and coordinates of NtcA-TAC and NtcA–NtcB-TAC have been deposited at the Electron Microscopy Data Bank (EMDB) (EMD-34476 for NtcA-TAC and EMD-34475 for NtcA–NtcB-TAC) and PDB (8H40 for NtcA-TAC and 8H3V for NtcA–NtcB-TAC). The supplementary map of NtcA–NtcB-TAC focusing on NtcA and NtcB has been deposited at EMDB under accession code of EMD-34477. Source data are provided with this paper.

References

43. Revyakin, A., Liu, C., Ebright, R. H. & Strick, T. R. Abortive initiation and productive initiation by RNA polymerase involve DNA scrunching. *Science* **314**, 1139–1143 (2006).
44. Kabsch, W. XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132 (2010).
45. Adams, P. D. et al. PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr. D Biol. Crystallogr.* **58**, 1948–1954 (2002).
46. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.* **53**, 240–255 (1997).
47. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).
48. Fang, C. L. et al. CueR activates transcription through a DNA distortion mechanism. *Nat. Chem. Biol.* **17**, 57–64 (2021).
49. Mekler, V., Pavlova, O. & Severinov, K. Interaction of *Escherichia coli* RNA polymerase σ^{70} subunit with promoter elements in the context of free σ^{70}, RNA polymerase holoenzyme, and the β′-σ^{70} complex. *J. Biol. Chem.* **286**, 270–279 (2011).
50. Mekler, V. & Severinov, K. RNA polymerase molecular beacon as tool for studies of RNA polymerase–promoter interactions. *Methods* **86**, 19–26 (2015).
51. Stark, H. GraFix: stabilization of fragile macromolecular complexes for single particle cryo-EM. *Methods Enzymol.* **481**, 109–126 (2010).
52. Chen, J., Noble, A. J., Kang, J. Y. & Darst, S. A. Eliminating effects of particle adsorption to the air/water interface in single-particle cryo-electron microscopy: bacterial RNA polymerase and CHAPSO. *J. Struct. Biol. X* **1**, 100005 (2019).
53. Zheng, S. Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).
54. Rohou, A. & Grigorieff, N. CTFFIND4: fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015).
55. Scheres, S. H. RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* **180**, 519–530 (2012).
56. Pettersen, E. F. et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).

Acknowledgements

We thank the staff at the Shanghai Synchrotron Radiation Facility (SSRF) for X-ray diffraction data collection of NtcB-EBD; P.-P. Tang and Y.-X. Gao at the Cryo-EM Center at University of Science and Technology of China for cryo-EM data acquisition. This work is supported by the Strategic Priority Research Program of the Chinese Academy of Sciences (http://www.cas.cn; grant nos. XDB37020301; C.-Z.Z., and XDA24020302; Y.-L.J.), the National Natural Science Foundation of China (http://www.nsfc.gov.cn; grant no. 32171198; Y.-L.J.), National Key Research and Development Program of China (grant no. 2018YFA0900700; Y.Z.) and Anhui Provincial Natural Science Foundation (http://kjt.ah.gov.cn; grant no. 2108085J14; Y.-L.J.). Y.-L.J. thanks the Youth Innovation Promotion Association of Chinese Academy of Sciences for their support (membership no. 2020452).

Author contributions

C.-Z.Z., Y.Z., Y.-L.J. and Y.C. conceived, designed and supervised the project. C.-Z.Z., Y.Z., Y.-L.J., Y.C. and S.-J.H. analyzed data and wrote the manuscript. S.-J.H., L.-Q.S., H.S., H.-C.M., K.Z. and N.C. performed the molecular cloning, protein expression and purification. S.-J.H., L.-L.Y. and W.-W.K. conducted the cryo-EM sample preparation and data acquisition. S.-J.H., Y.-L.J. and F.Y. performed cryo-EM data processing and model building. S.-J.H. and X.W. performed the biochemical assays. S.-J.H., Y.-L.J. and Z.-P.C. carried out the protein crystallization, X-ray data collection and structure determination. All authors discussed the data and read the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

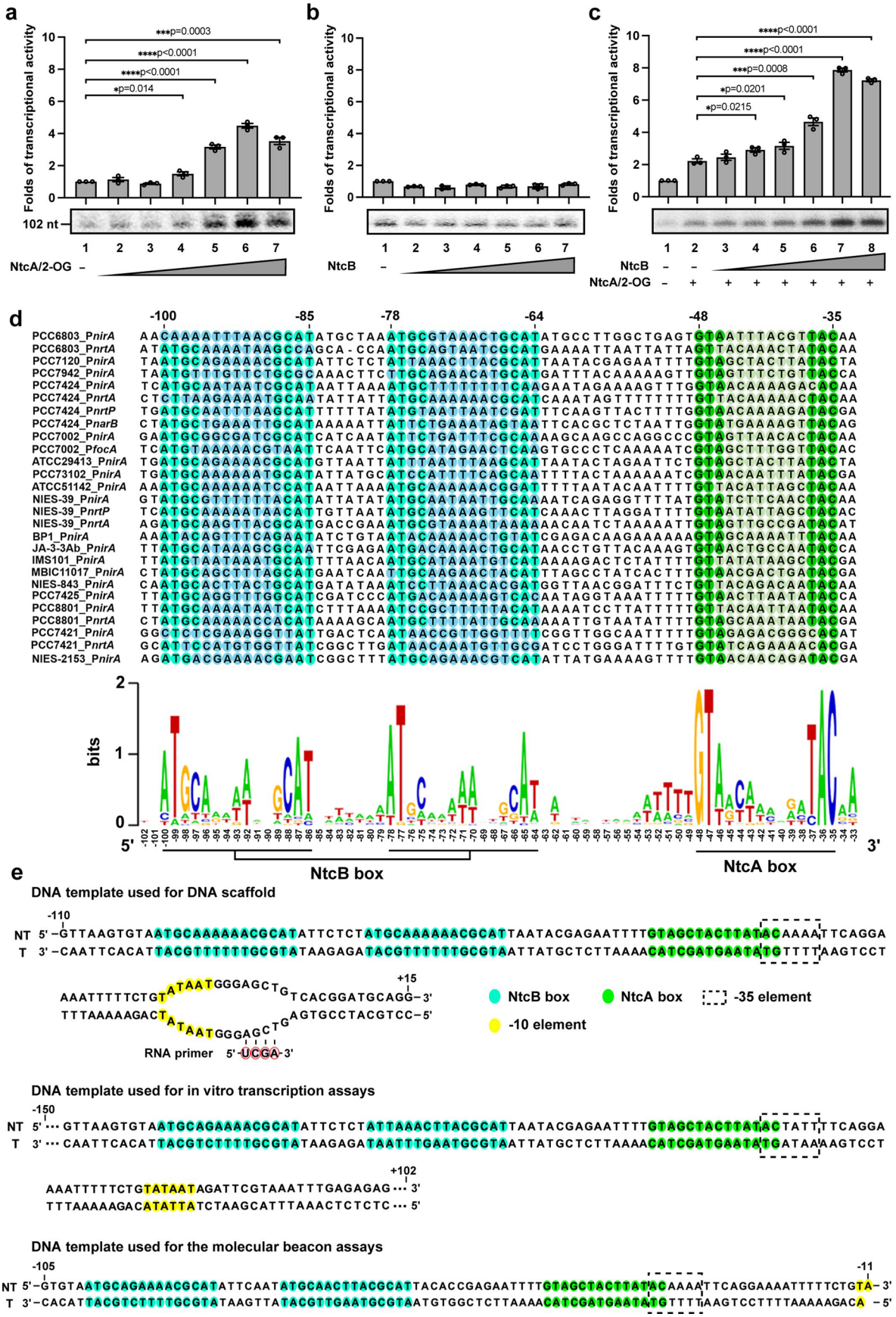
Extended data is available for this paper at https://doi.org/10.1038/s41594-023-01149-7.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41594-023-01149-7.

Correspondence and requests for materials should be addressed to Yong-Liang Jiang, Yu Zhang or Cong-Zhao Zhou.

Peer review information *Nature Structural & Molecular Biology* thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available. Primary Handling Editor: Sara Osman in collaboration with the *Nature Structural & Molecular Biology* team.

https://doi.org/10.1038/s41594-023-01149-7

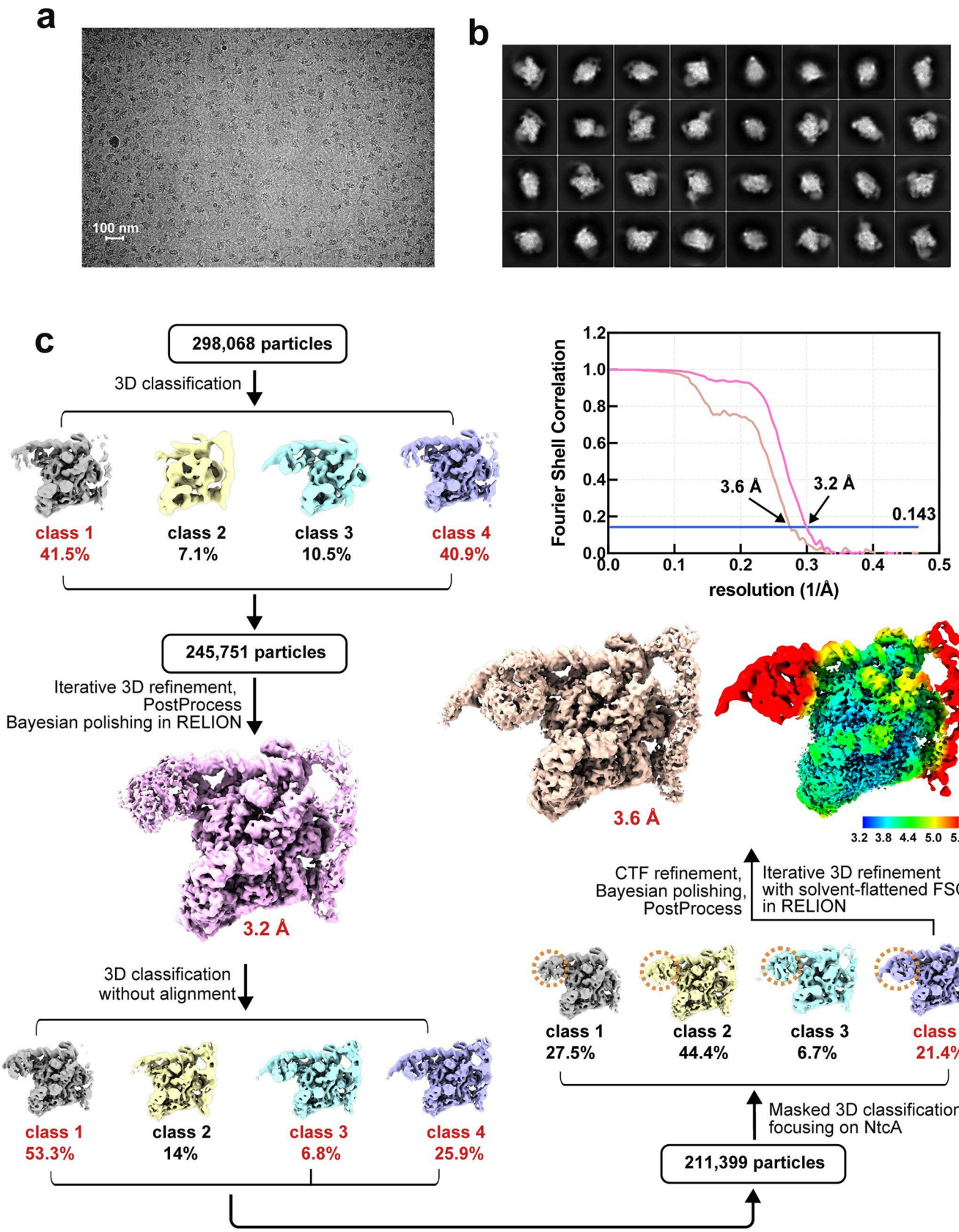


Extended Data Fig. 1 | See next page for caption.

https://doi.org/10.1038/s41594-023-01149-7

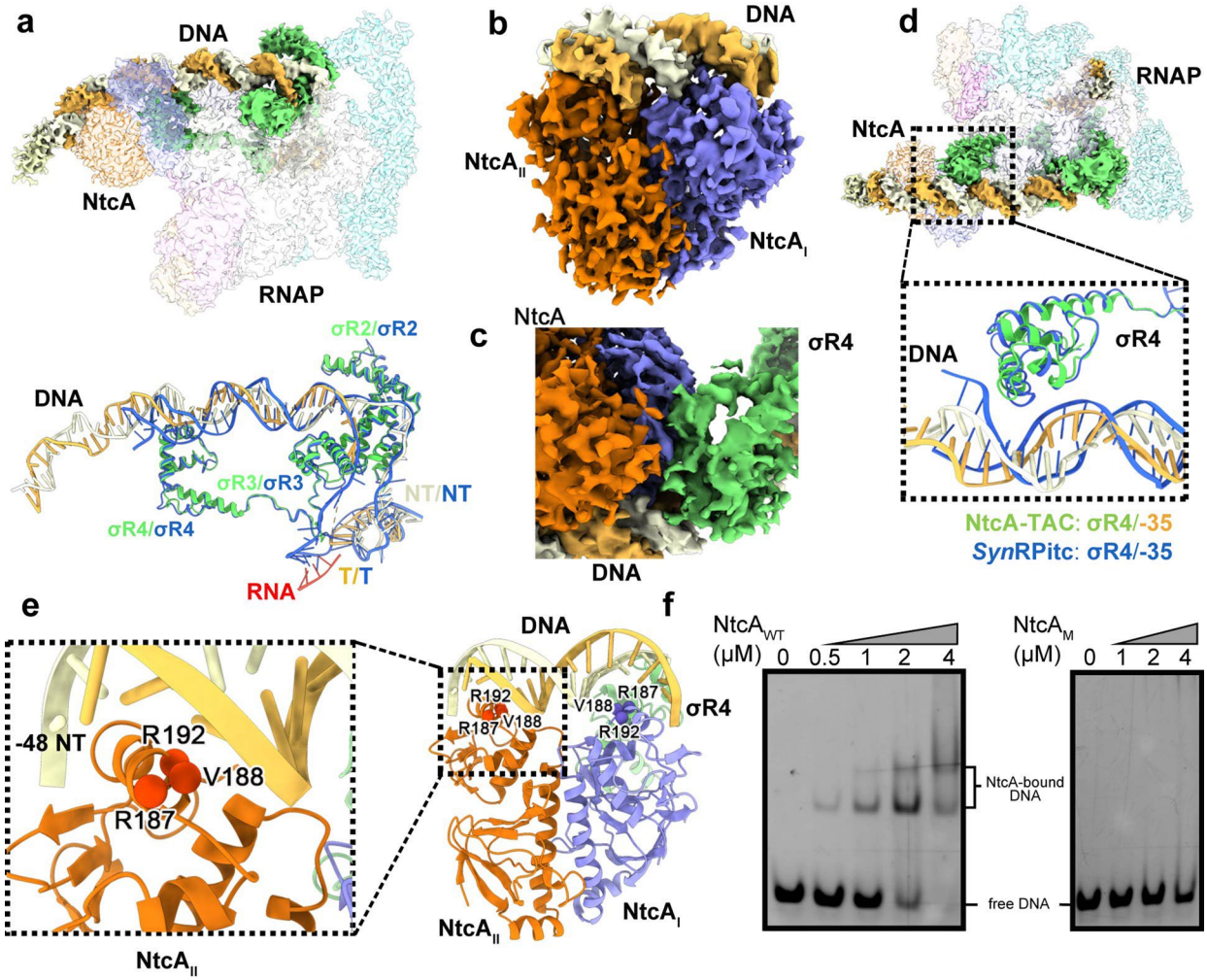
Extended Data Fig. 1 | The coordinated regulation of NtcA and NtcB on a PnirA derivative. a-b, The transcriptional activity of *Anabaena* sp. PCC 7120 RNAP in the presence of NtcA (**a**) or NtcB (**b**) at increasing concentrations (0, 25, 50, 100, 200, 400, and 800 nM). **c**, The transcriptional activity of RNAP preincubated with 200 nM NtcA, followed by adding NtcB at increasing concentrations (0, 50, 100, 200, 400, 800, and 1600 nM). Data are presented as mean ± S.E.M., n = 3 biologically independent experiments. Two-tailed unpaired t-tests. No p value adjustments for multiple comparisons. The lower panel shows the representative gel image. *p < 0.05, **p < 0.01, ***p < 0.001, and ****p < 0.0001 indicate significant difference compared to lane 1 (**a**) and lane 2 (**c**). RNA transcripts, 102 nt. **d**, The multiple-sequence alignment of the upstream promoter region of nitrate assimilation genes in β-cyanobacteria. The NtcA and NtcB boxes are colored green and cyan, respectively. The numbers on the top indicate the positions of NtcA and NtcB relative to transcription start site from *Anabaena* sp. PCC 7120. Sequence logos derived from the multiple-sequence alignment. **e**, The DNA templates used for DNA scaffold of Cryo-EM, *in vitro* transcription assays and the molecular beacon assays. NtcA and NtcB boxes, −35 element and −10 element are colored as shown in the color key. T, template; NT, nontemplate.

https://doi.org/10.1038/s41594-023-01149-7



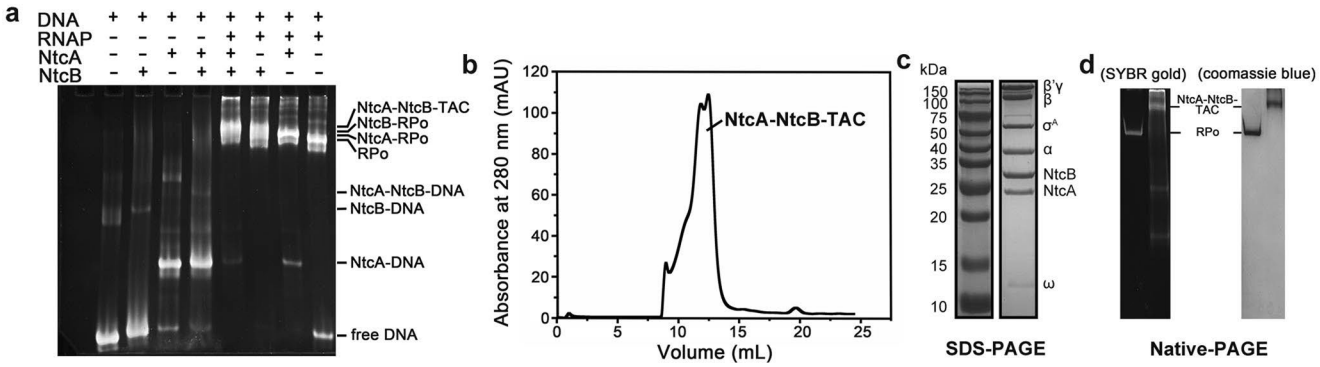
Extended Data Fig. 2 | Cryo-EM data processing of NtcA-TAC. a, A representative cryo-EM image of NtcA-TAC. Bar, 100 nm. The micrograph is a representative of 2,090 cryo-EM images. **b**, Representative 2D class averages of NtcA-TAC. **c**, Flowchart for cryo-EM data processing and map reconstruction for NtcA-TAC.

https://doi.org/10.1038/s41594-023-01149-7



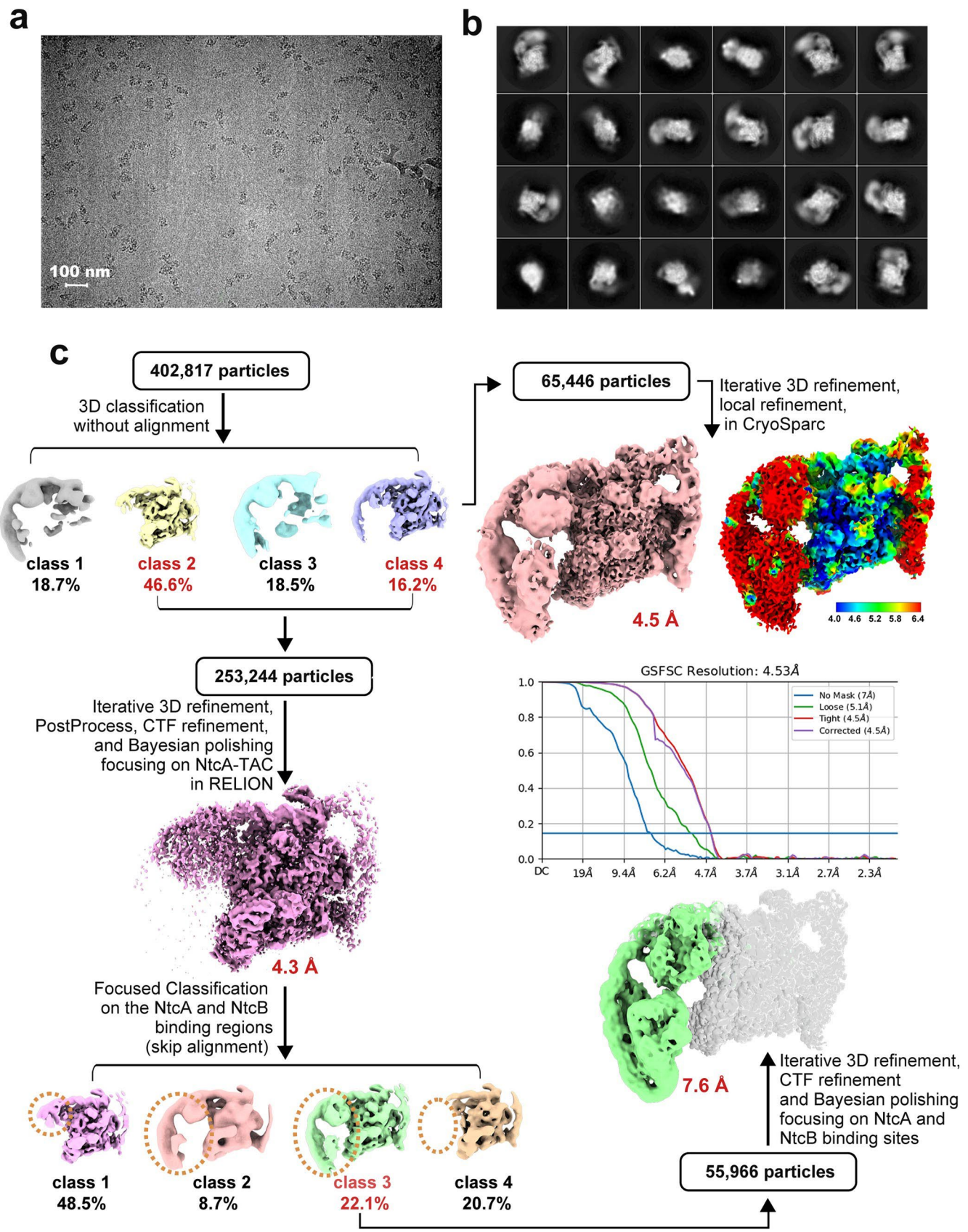
Extended Data Fig. 3 | The interaction pattern of NtcA binding to the promoter DNA and RNAP in NtcA-TAC. a, Superposition of promoter DNA between *Anabaena* NtcA-TAC and *Synechocystis* transcription initiation complex (*syn*RPitc, PDB: 8GZG). The color schemes are as follows: NT (nontemplate DNA) of NtcA-TAC, light yellow; T (template DNA) of NtcA-TAC, light orange; σ^A of NtcA-TAC, green; NT, T and σ^A of *syn*RPitc, blue; RNA of *syn*RPitc, red. The σR2, σR3, and σR4 represent the region 2, region 3 and region 4 of σ^A, respectively. **b**, Cryo-EM map of NtcA and its binding boxes. **c**, Cryo-EM maps of NtcA-σR4. **d**, Superposition of σR4-DNA between NtcA-TAC and *syn*RPitc. **e**, The interactions between NtcA and the NtcA box of promoter DNA. The key residues of NtcA involved in binding to DNA are shown as spheres. NT, nontemplate. **f**, The EMSA results showing binding of wild-type ($NtcA_{WT}$) or mutant (R187A/V188A/R192A; $NtcA_M$) NtcA with a DNA fragment containing the NtcA box. The EMSA analysis results is a representative (n = 1) of many times optimized experiments. Source data are provided in the Source Data file.

https://doi.org/10.1038/s41594-023-01149-7



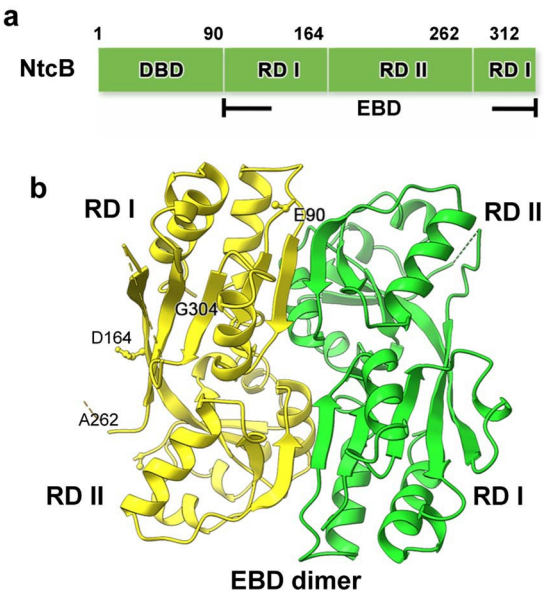
Extended Data Fig. 4 | In vitro reconstitution of NtcA-NtcB-TAC. a, Gel-shift result showing migration of protein-DNA complex reconstituted through indicated combinations of NtcA, NtcB, RNAP and the DNA in Fig. 1a. The Gel-shift result is a representative (n = 1) of many times optimized experiments. Source data are provided in the Source Data file. **b**, The size-exclusion chromatography of NtcA-NtcB-TAC. **c**, The SDS-PAGE analysis of the purified NtcA-NtcB-TAC complex. The protein components are labeled on the right of the gel. The chromatogram and SDS-PAGE are representatives of > 5 independent experiments that showed similar results. **d**, The native-PAGE analysis of NtcA-NtcB-TAC stained with SYBR Gold dye (left) or Coomassie Brilliant Blue (right). The native-PAGE analysis result is a representative of > 3 independent experiments that showed similar results. Source data are provided in the Source Data file.

https://doi.org/10.1038/s41594-023-01149-7



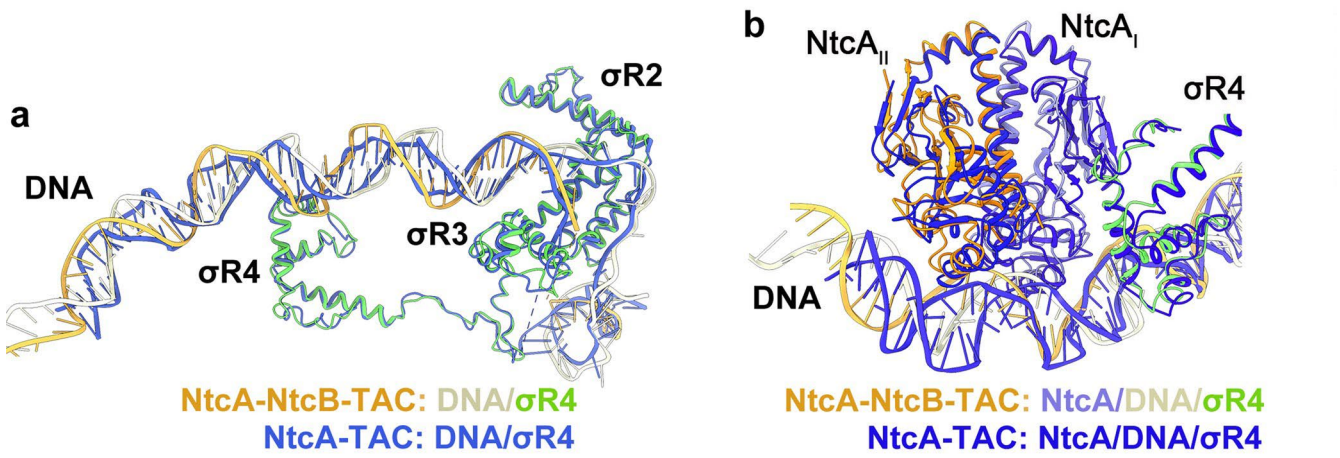
Extended Data Fig. 5 | Cryo-EM data processing of NtcA-NtcB-TAC. a, A representative cryo-EM image of NtcA-NtcB-TAC. Bar, 100 nm. The micrograph is a representative of 5,047 cryo-EM images. **b**, Representative 2D class averages of NtcA-NtcB-TAC. **c**, Flowchart for cryo-EM data processing of NtcA-NtcB-TAC.

https://doi.org/10.1038/s41594-023-01149-7



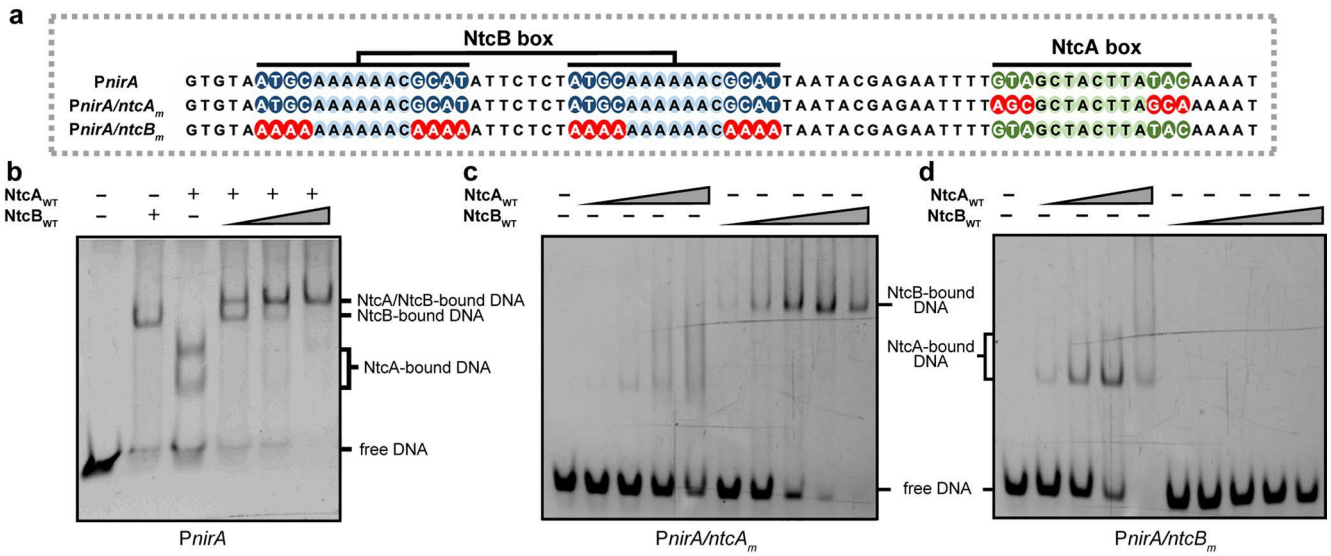
Extended Data Fig. 6 | The crystal structure of NtcB-EBD. a, A schematic presentation of domain organization of NtcB. **b**, Crystal structure of the NtcB-EBD dimer. The two subunits are colored yellow and green, respectively. The two subdomains RD I and RD II are labeled.

https://doi.org/10.1038/s41594-023-01149-7



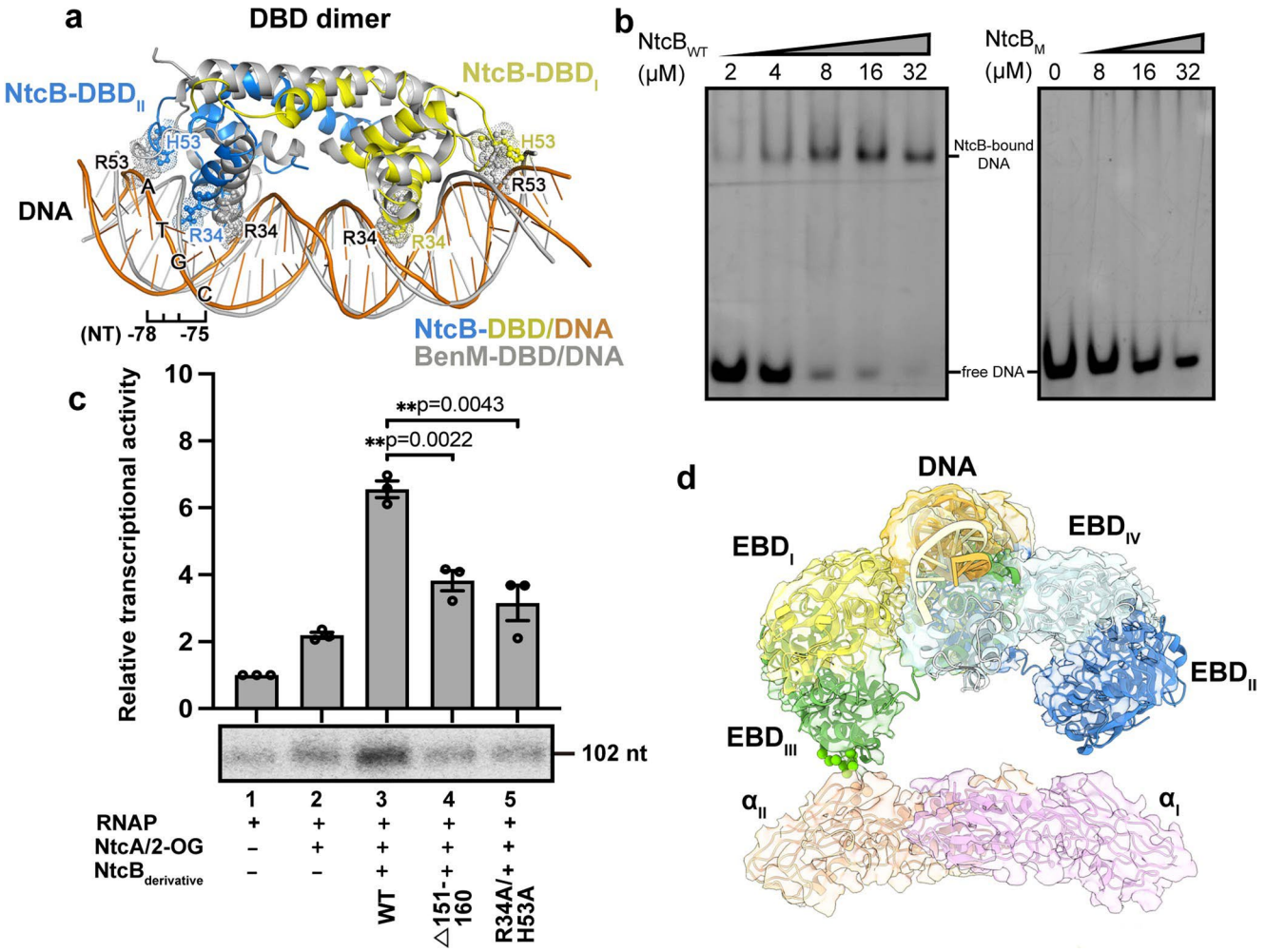
Extended Data Fig. 7 | Structural comparison of NtcA-TAC and NtcA-NtcB-TAC. a, Superposition of promoter DNA between NtcA-TAC and NtcA-NtcB-TAC. **b**, Superposition of the NtcA-σR4-DNA between NtcA-TAC and NtcA-NtcB-TAC.

https://doi.org/10.1038/s41594-023-01149-7



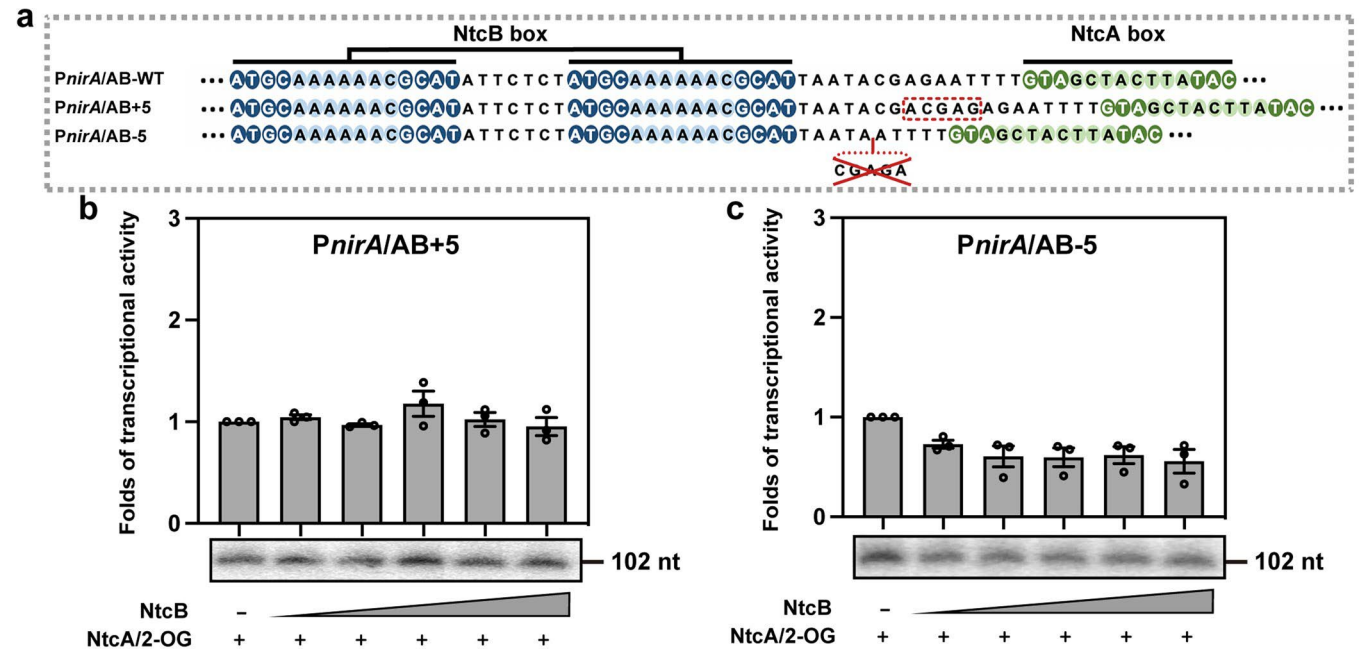
Extended Data Fig. 8 | NtcA and NtcB independently bind to their respective boxes. a, The sequences of 5′-Cy5 labeled dsDNA that are used in the EMSA assays. P*nirA*, 76-bp DNA containing wild-type NtcA and NtcB boxes; P*nirA/ntcA*$_m$, a P*nirA* derivative bearing mutations at the conserved palindromic nucleotides of NtcA box; P*nirA/ntcB*$_m$, a P*nirA* derivative bearing mutations at the conserved palindromic nucleotides of NtcB box. The mutated nucleotides are colored red. **b**, NtcA and NtcB bind to P*nirA* independently. NtcA$_{WT}$, wide-type NtcA protein; NtcB$_{WT}$, wide-type NtcB protein. **c**, Mutating the binding boxes of NtcA abolished NtcA binding activity but did not affect NtcB binding. **d**, Mutating the binding boxes of NtcB abolished NtcB binding activity but did not affect NtcA binding. All EMSA results are representatives of > 5 independent experiments that showed similar results. Source data are provided in the Source Data file.

https://doi.org/10.1038/s41594-023-01149-7



Extended Data Fig. 9 | The interactions of NtcB to promoter DNA and RNAP in NtcA-NtcB-TAC. a, Superposition of NtcB-DBD/DNA in NtcA-NtcB-TAC and the crystal structure of BenM-DBD/DNA complex (PDB:4IHS). The residues of NtcB (blue, yellow) and BenM (black) responsible for DNA motif recognition are labeled. NT, nontemplate. **b**, Alanine mutation of the DNA motif-recognition residues (R34A/H53A; NtcB$_M$) of NtcB abolished the binding ability of NtcB to a DNA fragment containing NtcB boxes. The EMSA results are representatives of > 3 independent experiments that showed similar results. Source data are provided in the Source Data file. **c**, Mutations of the DNA motif-recognition or RNAP-contact residues of NtcB impairs the transcription activation activity of NtcB. Data are presented as mean ± S.E.M., n = 3 biologically independent experiments. Two-tailed unpaired t-tests. No p value adjustments for multiple comparisons. The lower panel shows the representative gel image. **P < 0.05 indicates significant difference compared to lane 3. RNA transcripts, 102 nt. 2-OG, 2-oxoglutarate. △151–160, deletion of the contact patch (residues Leu151-Gly160) of NtcB. **d**, Cryo-EM map showing the interface between RNAP and NtcB-αNTD.

https://doi.org/10.1038/s41594-023-01149-7



Extended Data Fig. 10 | Alteration of spacer between NtcA- and NtcB-boxes abolishes transcription activation activity of NtcB. a, The sequences of three promoter DNAs that are used in *in vitro* transcription assays. P*nirA*/AB-WT, P*nirA* containing wild-type spacer length between NtcA and NtcB boxes; P*nirA*/AB + 5, a P*nirA* derivative containing insertion of the spacer between NtcA and NtcB boxes by 5 bp; P*nirA*/AB-5, a P*nirA* derivative containing deletion of the spacer between NtcA and NtcB boxes by 5 bp. The sequences of insertion or deletion of nucleotides are highlighted. **b**, **c**, Insertion (**b**) or deletion (**c**) of the spacer by 5 bp between NtcA and NtcB boxes impaired the transcription activation activity of NtcB. Data are presented as mean ± S.E.M., n = 3 biologically independent experiments. Two-tailed unpaired t-tests. No p value adjustments for multiple comparisons. The lower panel shows the representative gel image. RNA transcripts, 102 nt.

nature portfolio

Corresponding author(s): Cong-Zhao Zhou, Yu Zhang, Yong-Liang Jiang

Last updated by author(s): Aug 4, 2023

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐ A description of all covariates tested
☒	☐ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Cryo-EM images of NtcA-TAC and NtcA-NtcB-TAC were collected using EPU2 software and X-ray diffraction data of NtcB-EBD were collected at 100 K at beamline 19U with a DECTRIS PILATUS 6M detector. The radio graphs of in vitro transcription assay were obtained by storage-phosphor scanning (Typhoon; AMERSHAM TYPHOON version 5; GE Healthcare). The fluorescence emission intensities of the molecular beacon assays were recorded using a microplate reader (TECAN) and UJIFILM Multi Gauge (version 2.5).
Data analysis	Cryo-EM image analysis were performed using MotionCor2, CTFFIND 4, RELION 3.1, and cryoSPARC 3.1. The atomic models were built and refined by COOT 0.8.9.2 and PHENIX 1.18.2. All structures were validated by PHENIX 1.18.2 and MolProbity 4.02. ChimeraX 1.2.5 and Pymol 2.5.2 were used for preparing the structural figures, and in vitro transcription assays and molecular beacon assays data were analyzed with GraphPad Prism 9.1.1 and UJIFILM Multi Gauge 2.5. DNA sequence logos were generated by WebLogo (https://weblogo.threeplusone.com).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

The authors declare that the data supporting the findings of this study are available within the publication and its Supplementary Information have been deposited in the Protein Data Bank (PDB) or Electron Microscopy Data Bank (EMDB) as appropriate. The PDB accession code for NtcB-EBD, NtcA-TAC, and NtcA-NtcB-TAC were 8H3Z, 8H40, and 8H3V, respectively. The EMDB accession code of NtcA-TAC and NtcA-NtcB-TAC were EMD-34476 and EMD-34475, respectively. The supplementary map of NtcA-NtcB-TAC focusing on NtcA and NtcB has been deposited at EMDB under accession code of EMD-34477. Other PDB described in this study include 8GZG, 3LA2, and 4IHS.

Human research participants

Policy information about studies involving human research participants and Sex and Gender in Research.

Reporting on sex and gender	N/A
Population characteristics	N/A
Recruitment	N/A
Ethics oversight	N/A

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	The sample size has been described in the figure legends.
Data exclusions	No data was excluded.
Replication	The replicates for each experiments were described in the manuscript and figure legends.
Randomization	n/a (The current work doesn't involve population study)
Blinding	n/a (The current work doesn't involve experimental subjects with opinions)

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging