Supplementary Materials

Kinetics of transcription initiation directed by multiple *cis*-regulatory elements on the *glnAp2* promoter

Yaolai Wang[1], **Feng Liu**[1,2,†] and **Wei Wang**[1,2,‡]

[1]National Laboratory of Solid State Microstructures and Department of Physics, Nanjing University, Nanjing 210093, China;

[2]Collaborative Innovation Center of Advanced Microstructures, Nanjing University, Nanjing 210093, China.

[†‡]Correspondence and requests for materials should be addressed to F.L. (fliu@nju.edu.cn) or W.W. (wangwei@nju.edu.cn).

This document includes the following sections:

Supplemental Tables and Figures include Table S1-S3 and Figures S1-S10.

S1. Evaluation of possible DNA bridging mediated by NtrC

An NtrC hexamer has three DNA-binding domains (DBDs)[1]. The DBDs are connected to the oligomerization domains (ODs), which constitute the main ring of NtrC hexamer, by flexible polypeptide chains. Owing to the bending flexibility of DNA, two different binding sites in the regulatory region of the *glnAp2* promoter may be bridged by an NtrC hexamer under spatially and topologically favorable conditions. An NtrC tetramer, which exists during the formation of a hexamer, has two DBDs and thus may also bridge two different sites.

If any two of the five sites can be bridged, there exist three categories of bridging manners, i.e., enhancer–low-affinity site (including I-III, I-IV, I-V, II-III, II-IV, and II-V), low-affinity site–low-affinity site (including III-IV, III-V, and IV-V), and enhancer–enhancer (I-II) bridging. As detailed in the main text and Table S1, the two enhancers may only be transiently bridged by a tetramer.

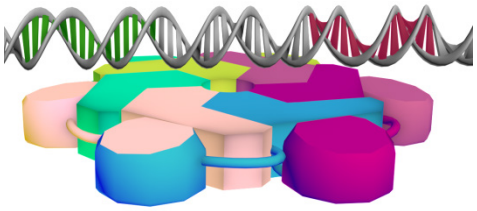
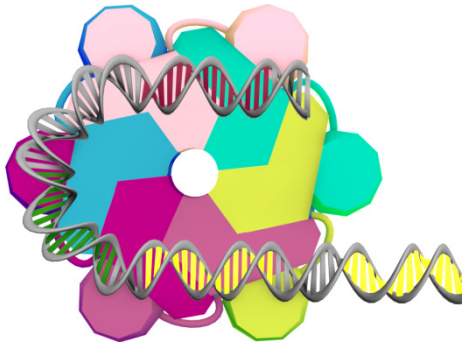
Notably, any bridging conformation of the first category — if it may occur — is unstable and the lifetime is rather short. A bridging by an NtrC hexamer involves four nodes, i.e., enhancer–DBD, OD_{enh}–main ring (OD_{enh} is the OD that is connected to the DBD bound to the enhancer), OD_{low}–main ring (OD_{low} is the OD that is connected to the DBD bound to the low-affinity site), and low-affinity site–DBD. A bridging by an NtrC tetramer has three nodes. The split of any node means breakdown of the bridging. Because of the instability of NtrC oligomer and low binding affinity of sites III-V for NtrC, all the bridging conformations are rather unstable. The bridging structures of the second category involve two low-affinity sites and one NtrC oligomer and thus are much more unstable.

To screen which bridging may occur, we explore each possible conformation in terms of its 3D structure, stability, dependence on the concentration of NtrC dimers, and potential influence on transcriptional output. By structure reconstruction, the spatial constraints such as the length of intervening DNA and the dimensions of NtrC hexamers are analyzed (Table S1). The formation of DNA bridging depends on the concentration of NtrC dimers. At low and intermediate concentrations, NtrC oligomers frequently form on the enhancers, whereas the low-affinity sites are rarely occupied. A hexamer/tetramer on either enhancer may approach a low-affinity site to induce DNA bridging of the first category. At high concentrations, the enhancers are occupied by hexamers most of the time, and the low-affinity sites are often occupied at least by NtrC dimers. The bridging conformations of the first category thus rarely occur, whereas those of the second category may occur by NtrC oligomerization.

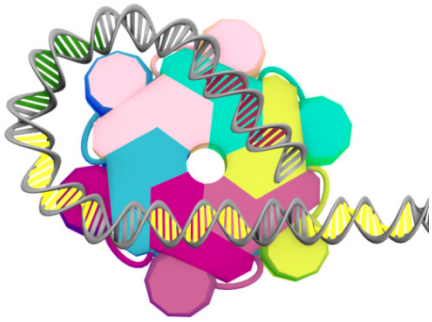
It is necessary to evaluate the degree to which each DNA bridging affects transcriptional dynamics and transcriptional output. This evaluation also facilitates overcoming the difficulty in determining the influence of topological factors, such as the degree of DNA torsion and the extent of a DBD detaching from and rotating around the main ring, on DNA bridging. The bridging configuration(s) that is/are topologically favored to occur must account for or at least be compatible with the complex transcriptional activities from both the wild-type and mutated *glnAp2* promoters. In conclusion, these detailed analyses reveal that the II-V bridging is the crucial conformation and plays a key role in transcriptional regulation.

Additional notes on the stability of DNA bridging:

According to the kinetic nature of key molecular interactions (see Table 1 in the main text), the average lifetime of an NtrC hexamer is ~4 min, while that of a DBD in association with an enhancer and a low-affinity site is 720 s and 72 s, respectively. Then, it can be inferred that, for an enhancer–low-affinity site bridging without involving serious DNA bending or torsion (like the II-V bridging), its average lifetime is ~55 s. For a low-affinity site–low-affinity site bridging without involving serious DNA bending or torsion, its average lifetime is ~33 s. With serious DNA bending or torsion involved, the lifetimes are much shorter.

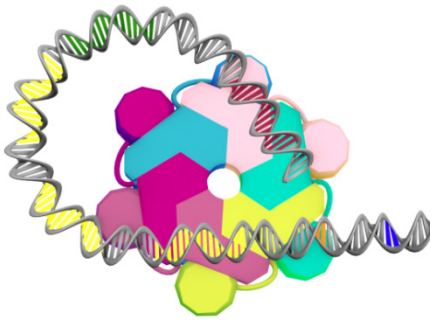
DNA bridging	Brief comment
I-II bridging 	The distance between two enhancers is much longer than that between two DBDs of an NtrC hexamer. The orientations of the two enhancers, toward which the DBDs insert into the major grooves, form an angle of ~37° (the periodicity of *in vivo* supercoiled DNA is ~11.1-11.2 bp per turn[2,3]). High energy is required to bend and even twist the short intervening DNA such that a hexamer can bridge enhancers I and II[1]. On the other hand, the putative cooperative binding of NtrC to the two enhancers is independent of the DBD and conformational change of DNA[4]. Therefore, the I-II bridging by hexamers never or hardly occurs. It remains controversial whether the binding of NtrC to the two enhancers exhibits cooperativity[5,6]. If the cooperativity exists, an eligible speculation would be that the oligomerization domain of a bound dimer helps recruit another dimer that is ultimately attracted to the other enhancer. In other words, the I-II bridging by a tetramer — if it transiently exists — slightly affects transcriptional dynamics by helping recruit NtrC dimers.
I-III bridging 	The I-III bridging hardly occurs. Obviously, when an NtrC tetramer or hexamer is formed at enhancer I, it is almost certain that enhancer II is also bound by an NtrC dimer, tetramer or hexamer, which hinders DNA looping.

I-IV bridging



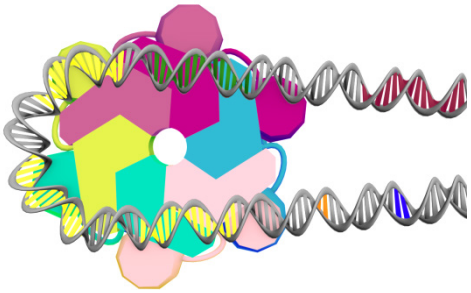
Similar to the I-III bridging, the formation of I-IV bridging is also inhibited due to enhancer II, although the intervening DNA is much longer. If this conformation occurs occasionally, it represses transcription by preventing NtrC hexamers at any site from contacting the holoenzyme.

I-V bridging



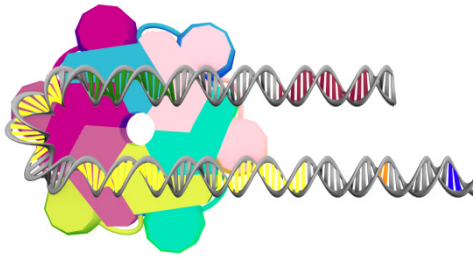
Similar to the I-IV bridging, the I-V bridging inhibits transcription by preventing hexamers at any site from contacting the holoenzyme if it might appear.

II-V bridging



Sites II and V can be bridged by an NtrC hexamer or tetramer. Given the low-affinity sites are rarely occupied at low and intermediate concentrations of NtrC dimers, the II-V bridging forms when an enhancer II-bound NtrC oligomer encounters site V. The resulting II-V bridging exactly constrains enhancer I in the vicinity of the -12 region, facilitating the hexamer at enhancer I to rapidly find and catalyze the holoenzyme. At high concentrations, the II-V bridging rarely forms because of the occupancy of sites III and IV. The II-V bridging thus accounts for the contributions of low-affinity sites to elevated transcriptional output at low and intermediate NtrC concentrations.

II-IV bridging



If the II-IV bridging might occur, it represses transcription by rendering hexamers at enhancer I hard to approach the -24~-12 region. Moreover, the II-IV bridging competes with the II-V bridging for enhancer II-bound NtrC hexamer/tetramer.

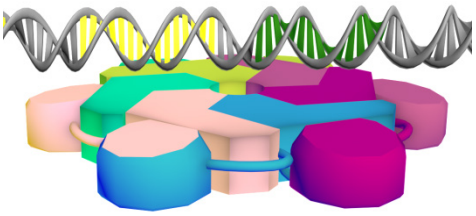
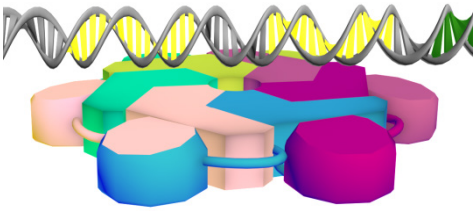
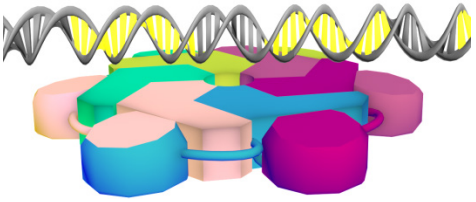
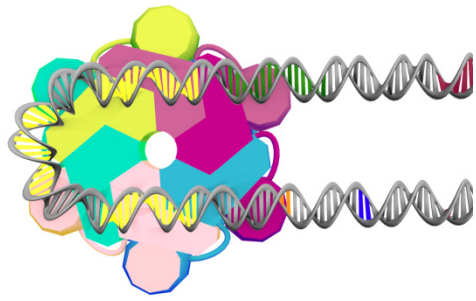
II-III bridging 	The orientations of sites II and III point to two directions with an angle of ~60°. Sites II and III are thus unlikely to be bridged by an NtrC hexamer. Even if this conformation appears, the hexamer lacks the degree of freedom to contact the holoenzyme, compared with an enhancer II-bound hexamer that can stimulate transcription initiation. If sites II and III are bridged by a tetramer occasionally, this conformation is rather unstable and does not affect transcriptional dynamics.
III-IV bridging 	The III-IV bridging may occur only at very high concentrations. The hexamer bridging sites III and IV cannot contact the holoenzyme because site V is also occupied. Additionally, this configuration hinders the enhancer-bound hexamers from approaching the holoenzyme; this effect is equivalent to that due to the occupancy of sites III and IV by individual NtrC dimers or oligomers.
IV-V bridging 	The IV-V bridging may occur only at very high concentrations. If it occurs, the hexamer cannot contact the holoenzyme to activate transcription, and this effect is equivalent to transcriptional inhibition by the occupancy of sites IV and V.
III-V bridging 	The III-V bridging hardly occurs. When sites III and V are occupied, site IV is also bound by an NtrC dimer, tetramer, or hexamer, which hinders DNA looping.

Table S1. Evaluation of possible DNA bridging mediated by NtrC. The five sites and -24~-12 region are colored in the same way as in Figure 1A of the main text. For the location of the DBD, refer to Figure 1B.

S2. Timescale of DNA looping

There exist three crucial modes of DNA looping, i.e., II-V bridging, and an enhancer I/II-bound NtrC hexamer contacting the closed complex. The shortest DNA loop in this study is the II-V bridging, involving 63 base pairs and no serious bending. To bend such a DNA segment, the main obstacle is conformational entropy[7-9].

After two ends of a DNA segment with n base pairs are connected, the change in entropy is defined as

$$\Delta S = -k_B \, ln \frac{\Omega_{DNA_{open}}}{\Omega_{DNA_{loop}}}, \tag{1}$$

where k_B is the Boltzmann constant, $\Omega_{DNA_{open}}$ and $\Omega_{DNA_{loop}}$ are separately the number of all possible conformations of the DNA segment in "open" and "looping" states. The corresponding change in the free energy, ΔG, satisfies

$$\Delta G = -T\Delta S, \tag{2}$$

where T is the absolute temperature. The average time for the DNA segment to convert from an open state to a looping state, τ, obeys[10]

$$\tau^{-1} = \frac{k_B T}{h} \, e^{-\Delta G/k_B T}, \tag{3}$$

where h is Plank's constant. For two DNA segments with n and n_0 base pairs, τ and τ_0 satisfy

$$ln\frac{\tau}{\tau_0} = \frac{1}{k_B T}(\Delta G - \Delta G_0) = -\frac{1}{k_B T}(T\Delta S - T\Delta S_0).$$

That is,

$$ln\frac{\tau}{\tau_0} = -\frac{1}{k_B}(\Delta S - \Delta S_0). \tag{4}$$

According to Jacobson & Stockmayer function[8,11],

$$\Delta S - \Delta S_0 = -\, c \, k_B \ln\left(\frac{n}{n_0}\right), \tag{5}$$

with $c = 1.75$. Now we have

$$ln\frac{\tau}{\tau_0} = c \, \ln\left(\frac{n}{n_0}\right),$$

i.e.,

$$\frac{\tau}{\tau_0} = \left(\frac{n}{n_0}\right)^{1.75}. \tag{6}$$

The two ends of a DNA segment are not directly connected, but are mediated by a protein at one end (or two proteins at both ends). A specific binding site at the surface of this bound protein associates with

the other binding site at the other end (or at the protein therein).The configuration number of such mediator proteins is not considered in the derivation above. Taking this into account, the change in entropy obeys

$$\Delta S = -k_B \, ln \frac{\Omega'_{DNA_{open}} \, \Omega_{protein}}{\Omega'_{DNA_{loop}} \, \Omega_{protein_{connected}}} . \tag{7}$$

Since the configuration number of the DNA is restrained by the protein(s), $\Omega'_{DNA_{open}} < \Omega_{DNA_{open}}$ and $\Omega'_{DNA_{loop}} < \Omega_{DNA_{loop}}$. Let $\Omega_{DNA_{open}} = \alpha \, \Omega'_{DNA_{open}}$ and $\Omega_{DNA_{loop}} = \beta \, \Omega'_{DNA_{loop}}$, where α ($\alpha > 1$) and β ($\beta > 1$) are two constants. Considering the restraint on configuration number by the directivity of protein-DNA or protein-protein interactions at the two ends, $\alpha < \beta$. Then,

$$\Delta S = -k_B \ln \left(\frac{\Omega_{DNA_{open}}}{\Omega_{DNA_{loop}}} \, \frac{\beta}{\alpha} \, \frac{\Omega_{protein}}{\Omega_{protein_{connected}}} \right),$$

$$\Delta S = -k_B \, ln \frac{\Omega_{DNA_{open}}}{\Omega_{DNA_{loop}} f} , \tag{8}$$

where f is a constant, i.e., $f = \frac{\alpha}{\beta} \frac{\Omega_{protein_{connected}}}{\Omega_{protein}}$. Since $\Omega_{protein} > \Omega_{protein_{connected}}$, $0 < f < 1$. f depends on the change in protein configuration number, as well as on the change in DNA configuration number due to the directivity of protein-DNA or protein-protein interactions at the two ends. In other words, f is affected by topological factors such as the degree of DNA torsion. When a protein at one end of the DNA segment approaches the other end (or the protein therein), DNA looping forms if the two binding sites face each other. If facing each other requires a tenser DNA torsion, which represents a more ordered thermodynamic state, f should take a smaller value, and the probability of forming the looping is lower. That is, f can be considered as the probability of connecting two sites to form a loop.

Referring to Eq. (5), we have

$$\Delta S - \Delta S_0 = -c \, k_B \ln \left(\frac{n}{n_0} \right) - k_B \ln \frac{1}{f}, \tag{9}$$

Therefore,

$$\frac{\tau}{\tau_0} = \frac{1}{f} (\frac{n}{n_0})^c. \tag{10}$$

For $n_0 = 468$ bp, it was experimentally shown that $\tau_0 = 526$ s[12]. Thus,

$$\tau = 0.0112 \, \frac{n^{1.75}}{f} . \tag{11}$$

Notably, these formulas are based on the condition that the intervening DNA is unoccupied by any

protein.

Our structural analyses have revealed that, among possible architectures of the transcription apparatus, three configurations are responsible for transcription initiation. That is, an NtrC hexamer at enhancer I or II catalyzes the holoenzyme via DNA looping, and the II-V bridging facilitates the enhancer I-mediated transcription initiation. These configurations are topologically favored to occur, thereby allowing for effective transcriptional regulation. Moreover, it can be inferred that the enhancer II-driven transcription initiation is less prominent for two reasons. 1) The orientations of the two enhancers, toward which a DBD inserts into the major groove of DNA, point to two directions with an angle of ~37°. Thus, the active centers of hexamers at the two enhancers face different directions, and their probabilities for contacting the catalysis site of the holoenzyme are different due to the rigidity in DNA torsion[9,13]. This implies that either enhancer I or II-bound hexamers are less favored to stimulate transcription. 2) The enhancer II-bound hexamers are engaged in either contacting the holoenzyme or bridging sites II and V; the II-V bridging should take priority, since it is the crucial configuration that underlies the observed transcriptional activities.

Taken together, the structural basis for transcriptional regulation suggests that, when a hexamer at enhancer I encounters the holoenzyme, its active center should precisely contact the catalysis site of the holoenzyme. That is, for enhancer I-bound hexamers to contact the holoenzyme, f_{I-H} is close to 1. A free DBD of a hexamer bound to enhancer II can smoothly find site V to induce the II-V bridging, whereas the DBD is less likely to contact the holoenzyme. That is, f_{II-V} is close to 1 but f_{II-H} is to be further determined. According to Eq. (11), it takes 16 s to form the II-V bridging and 64 s for a hexamer at enhancer I to contact the holoenzyme on average. In numerical simulations, we found that these two values indeed enable a perfect reproduction of various data on transcriptional activities. We determined f_{II-H} by fitting to the experimental data (the details of fitting are available in Supplemental S3) and got $f_{II-H} \approx 0.5$. That is, the average time required for an enhancer II-bound hexamer to contact the holoenzyme is 80 s.

S3. Stochastic model for NtrC-regulated *glnAp2* transcription

The dynamics of NtrC-regulated *glnAp2* transcription can be detailed by seventy reactions (Figure S1 and Table S2). Reactions 1-45 include five groups of reactions, with each describing the dynamics of NtrC interacting with one of the five sites. Reactions 46-60 describe the formation and disaggregation of the II-V bridging. Reactions 61-70 describe the kinetics of σ^{54}RNAP recognizing the promoter DNA and transcription initiation directed by NtrC hexamers.

The kinetics of the σ^{54}RNAP interacting with promoter DNA were measured on the *glnALG* promoter[12,14]. These data can be safely applied to the *glnAp2* promoter because these two promoters share the same sequences in the -24~-12 and neighboring regions. Owing to the simplicity of the transcription apparatus that only involves NtrC, σ^{54}RNAP and promoter DNA, only few parameters need be determined (see Table S2). On the other hand, abundant quantitative data on transcriptional dynamics are available[5,9,15]. At low and intermediate concentrations of NtrC dimers, there are seven data points on the transcriptional input-output curve in Fig. 6 of Ref. 5. At high concentrations, there are five data points in Fig. 3 of Ref. 15. From these data, it can be inferred that half the maximal initiation rate occurs at ~2 nM and ~400 nM. Those parameter values thus can be easily determined based on physical laws and numerical fitting to these twelve data points. Robustness analyses were then made to justify the estimation.

The fitting method is as follows. For a given NtrC concentration, we simulated the evolution of molecular reactions using the Gillespie algorithm[16,17]. After discarding the transients that depend on initial values of reactants, the number of mRNA produced in an interval of 5 min was sampled. With a sufficiently large sample number of 40 000, we obtained the average rate of mRNA production. By repeating the above process for every concentration, we obtained the transcriptional input-output function (I-O function). The I-O function was then normalized by setting the maximum to 1.0. We then compared this I-O function with that determined by the twelve experimental data points (shown in Figure 4A). We continued to optimize the parameter values until the best global consistency was achieved using the method of least squares.

In the following are introductions to the 70 reactions. Reactions 1-9, 10-18, 19-27, 28-36, and 37-45 separately describe the state evolution of the five binding sites. An NtrC dimer binds to an enhancer or a low-affinity site and then nucleates other dimers to form a tetramer and then a hexamer. The tetramers and

hexamers are unstable and disaggregate. These processes involve three types of molecular reactions including DBD–enhancer, DBD–low-affinity site, and oligomerization domain–oligomerization domain interactions.

Let τ_e and τ_l separately denote the average duration of a DBD bound to an enhancer and to a low-affinity site. The corresponding dissociation rates are separately $1/\tau_e$ and $1/\tau_l$. For a hexamer X composed of three dimers A, B, and C, its disaggregation is expressed as follows: $X \xrightarrow{k_o} AB + C$, $X \xrightarrow{k_o} AC + B$, or $X \xrightarrow{k_o} BC + A$, where $k_o$is the reaction rate. The average lifetime of a hexamer is thus $1/(3k_o)$. Let $\tau_o = 1/k_o$; τ_o is the average duration of a dimer in a hexamer. The disaggregation of a tetramer AB is expressed as $AB \xrightarrow{k_o'} A + B$. Simply, $k_o' = k_o$ (this approximation is sufficient here; see endnote of this section①). The fitting reveals that τ_e =720 s, τ_l= 72 s, and τ_o=720 s. Accordingly, the average lifetime of NtrC hexamer is 240 s.

To describe the association reactions, we employed the concept of "median number" of NtrC dimers (see Ref. 18 for details of this method). Let m_e, m_l, and m_o separately denote the median number when describing NtrC dimers binding to an enhancer, a low-affinity site, or an NtrC dimer/tetramer. Take the definition of m_e as an example. When the number of NtrC dimers, n, equals m_e, the rates of association and dissociation between an enhancer and NtrC dimers are identical. The ratio n/m_e is the effective number of NtrC dimers; in simulations, n/m_e ranges from 0.01 to 10000, covering the total concentration range to which the *glnAp2* promoter may respond. The fitting reveals that m_e: m_l: m_o =2:200:1. Since half the maximal initiation rate first occurs at ~2 nM[4,5], we have m_e =2 nM, m_l =200 nM, and m_o =1 nM.

Reactions 46-52 and 53-60 separately describe the formation and disaggregation of II-V bridging. The rate constant for forming the II-V bridging is 1/16 s^{-1} when neither of sites III and IV is occupied. This value was calculated based on Eq. (11) in S2. When site III or IV is occupied, the rate constant is 1/160 s^{-1}. When sites III and IV are occupied, the rate constant is 0, i.e., the II-V bridging never forms (Figure S5). These two values were obtained by fitting to the experimental data on the production rate at high NtrC concentrations[15].

Reactions 61-64 and 69-70 describe the dynamics of the holoenzyme recognizing the core promoter and initiating transcription. The number of σ^{54}RNAP is assumed to be sufficient and constant; in simulations, we used 0.03 s as the time for σ^{54}RNAP to approach its binding site. Experimentally, the

upper limit of this time is no more than 0.08 s[14]. Since the absolute value of maximal transcription rate is unavailable experimentally, we used a relative value as the output (i.e., the transcription rate is normalized by setting the maximum to 1.0). Therefore, our results are independent of the searching time.

Reactions 65-68 describe the dynamics of enhancer-bound hexamers getting to contact the holoenzyme via looping the intervening DNA. The time for DNA looping is based on theoretical calculations, as detailed in Supplemental S2. The time required for NtrC hexamers at enhancer II to contact the holoenzyme is optimized to be 80 s. With this value, *glnAp2* is transcribed at 45% of the wild-type level, exactly consistent with the experimental data. In contrast, halving or doubling this value would make the transcription rate in Case 1 (III-IV-V-) be 59% and 39% of the wild-type level, respectively. The looping time takes different values depending on the status of the intervening DNA. To get a good fitting to the data on the initiation rate at high NtrC concentrations[15], enhancer I/II-bound hexamers should not contact the holoenzyme when the three low-affinity sites are all occupied, or reach the holoenzyme with a small probability when one or two sites are occupied (Figure S6). The average time for a hexamer at enhancer I to approach the holoenzyme is estimated at 3 s given the II-V bridging. This estimate is based on the consideration that hexamers at enhancer I are just near the -12 region when the II-V bridging is formed. In fact, this value does not affect our conclusions; halving or doubling it leads to a slight change of transcriptional behavior (Figure S7). The average time for an enhancer-bound hexamer to catalyze the holoenzyme is estimated at 1 s based on the timescale of biochemical reactions; robustness analysis shows that even halving or doubling this value also does not affect the transcriptional behaviors (Figure S8).

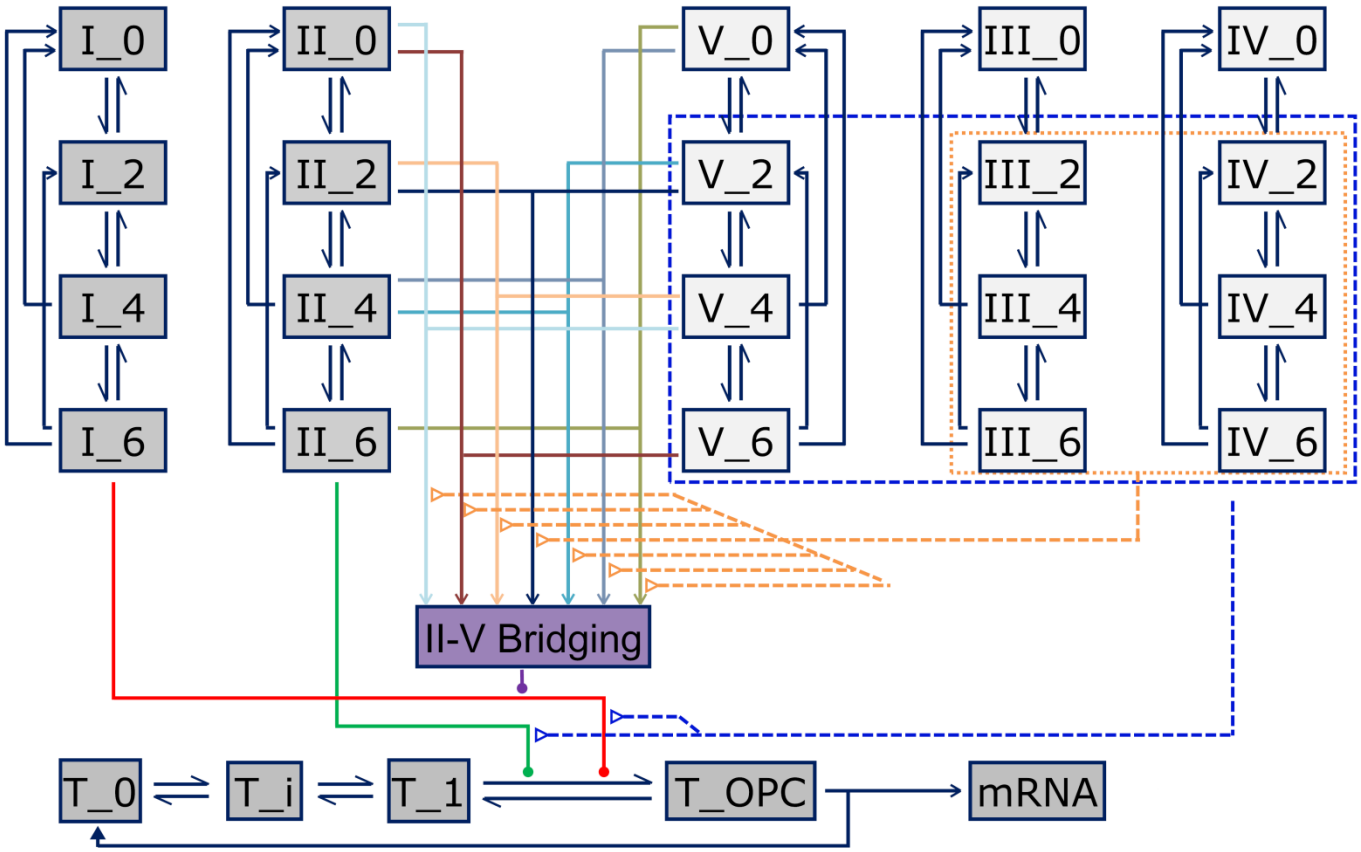


Figure S1. Detailed stochastic model of NtrC-regulated *glnAp2* expression. The state of an enhancer (I or II) or a low-affinity site (III, IV, or V), converts stochastically among being vacant, bound by an NtrC dimer, tetramer or hexamer (separately denoted by suffixes "0", "2", "4" and "6"). On the core promoter ("T_0") where the holoenzyme binds, two closed complexes ("T_i" and "T_1") are sequentially formed. The posterior closed complex can isomerize into the open complex ("T_OPC"), once catalyzed by an NtrC hexamer bound to enhancer I or II (separately denoted by red and green lines ending with blunt arrows). Enhancer II and site V can be bridged by an NtrC tetramer or hexamer (different reaction pathways are distinguished by different colors). The II-V bridging facilitates the enhancer I-bound hexamers to stimulate transcription initiation (denoted by a violet line ending with a blunt arrow). Site III or IV bound by an NtrC dimer/tetramer/hexamer hinders the formation of II-V bridging (denoted by an orange box and orange lines ending with hollow triangles). Any low-affinity site bound by NtrC represses transcription initiation by blocking the enhancer-bound hexamers from approaching the holoenzyme (denoted by a blue box and blue lines ending with hollow triangles).

	Reaction	rate constant (s^{-1})
1	I_0+ N_d → I_2	$n/(\tau_e m_e)$
2	I_2 → I_0	$1/\tau_e$
3	I_2 + N_d → I_4	$n/(\tau_o m_o)$
4	I_4 → I_2	$1/\tau_o$
5	I_4 + N_d → I_6	$n/(\tau_o m_o)$
6	I_6 → I_4	$2/\tau_o$
7	I_6 → I_0	$1/\tau_e$
8	I_4 → I_0	$1/\tau_e$
9	I_6 → I_2	$1/\tau_o$
10	II_0 + N_d → II_2	$n/(\tau_e m_e)$
11	II_2 → II_0	$1/\tau_e$
12	II_2 + N_d → II_4	$n/(\tau_o m_o)$
13	II_4 → II_2	$1/\tau_o$
14	II_4 + N_d → II_6	$n/(\tau_o m_o)$
15	II_6 → II_4	$2/\tau_o$
16	II_6 → II_0	$1/\tau_e$
17	II_4 → II_0	$1/\tau_e$
18	II_6 → II_2	$1/\tau_o$
19	III_0 + N_d→ III_2	$n/(\tau_l m_l)$
20	III_2 → III_0	$1/\tau_l$
21	III_2 + N_d → III_4	$n/(\tau_o m_o)$
22	III_4 → III_2	$1/\tau_o$
23	III_4 + N_d → III_6	$n/(\tau_o m_o)$
24	III_6 → III_4	$2/\tau_o$
25	III_6 → III_0	$1/\tau_l$
26	III_4 → III_0	$1/\tau_l$
27	III_6 → III_2	$1/\tau_o$
28	IV + N_d→ IV_2	$n/(\tau_l m_l)$
29	IV_2 → IV_0	$1/\tau_l$
30	IV_2 + N_d → IV_4	$n/(\tau_o m_o)$
31	IV_4 → IV_2	$1/\tau_o$
32	IV_4 + N_d → IV_6	$n/(\tau_o m_o)$
33	IV_6 → IV_4	$2/\tau_o$
34	IV_6 → IV_0	$1/\tau_l$
35	IV_4 → IV_0	$1/\tau_l$
36	IV_6 → IV_2	$1/\tau_o$
37	V_0 + N_d → V_2	$n/(\tau_l m_l)$
38	V_2 → V_0	$1/\tau_l$
39	V_2 + N_d → V_4	$n/(\tau_o m_o)$
40	V_4 → V_2	$1/\tau_o$
41	V_4 + N_d → V_6	$n/(\tau_o m_o)$
42	V_6 → V_4	$2/\tau_o$
43	V_6 → V_0	$1/\tau_l$
44	V_4 → V_0	$1/\tau_l$

Parameters

Reactions 1-45 include five groups of reactions, with each describing the kinetics of NtrC interacting with one of the five sites. The rate constants of these reactions only involve four independent parameters:

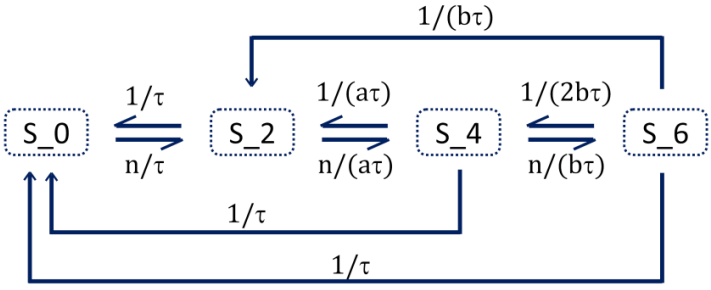
τ_e =720 s;
τ_l = 72 s;
τ_o =720 s;
m_e: m_l: m_o = 2:200:1.

45	V_6 → V_2	$1/\tau_o$	
46	II_0+V_6→II-V_6	$1/\tau_{25}$	Reactions 46-60 describe the kinetics of the II-V bridging. $1/\tau_{25}$ = 1/16 s^{-1}, 1/160 s^{-1}, and 0 s^{-1} —separately corresponding to the cases where neither of, either of, and both of sites III and IV is/are bound by NtrC.
47	II_0+V_4→II-V_4		
48	II_2+V_2→II-V_4		
49	II_2+V_4→II-V_6		
50	II_4+V_0→II-V_4		
51	II_4+V_2→II-V_6		
52	II_6+V_0→II-V_6		
53	II-V_4+N_d→II-V_6	$n/(\tau_o m_o)$	
54	II-V_4→II_2+V_2	$1/\tau_o$	
55	II-V_4→II_4+V_0	$1/\tau_l$	
56	II-V_4→II_0+V_4	$1/\tau_e$	
57	II-V_6→II_6+V_0	$1/\tau_l$	
58	II-V_6→II_4+V_2	$1/\tau_o$	
59	II-V_6→II_2+V_4	$1/\tau_o$	
60	II-V_6→II_0+V_6	$1/\tau_e$	
61	T_0+σ^{54}RNAP→T_i	*1/0.03;*	Reactions 61-70 describe the kinetics of transcription initiation. $1/\tau_{IT}$ = 1/3 s^{-1}, 1/64 s^{-1}, 1/128 s^{-1}, 1/640 s^{-1},and 0 s^{-1} — separately corresponding to the cases where the II-V bridging forms, none of, one of, two of, and all of sites III-V is/are bound by NtrC. $1/\tau_{IIT}$ = 1/80 s^{-1}, 1/160 s^{-1}, 1/640 s^{-1}, and 0 s^{-1} — separately corresponding to the cases where none of, one of, two of, and all of sites III-V is/are bound by NtrC.
62	T_i→T_0	*1/3*	
63	T_i→T_1	*1/10*	
64	T_1→T_i	*1/125*	
65	I_6+T_1→I-T	$1/\tau_{IT}$	
66	II_6+T_1→II-T	$1/\tau_{IIT}$	
67	I-T→I_6+T_OPC	*1/1*	
68	II-T→II_6+T_OPC		
69	T_OPC →T1	*1/526*	
70	T_OPC→T0+mRNA	*1/5.8*	

Table S2. Reactions and reaction rates in the stochastic model. I_0, I_2, I_4, and I_6 separately denote that enhancer I is unbound, bound by an NtrC dimer, tetramer, and hexamer; and so are the other four sites II-V. N_d denotes one free NtrC dimer and n denotes the number of NtrC dimers. τ_e and τ_l separately denote the average duration of a DBD in association with an enhancer and a low-affinity site. τ_o denotes the average time of an NtrC dimer in association with an NtrC dimer or tetramer. n/m_e, n/m_l, and n/m_o are separately the effective number of NtrC dimers when computing the reaction rates of a dimer binding to an enhancer, a low-affinity site, or a bound dimer/tetramer. S/S_m is the effective number of σ^{54}RNAP when computing the reaction rates of σ^{54}RNAP binding to the core promoter.

①Endnote: More detailed description of NtrC interacting with its binding site.

A much more detailed description of NtrC interacting with its binding site is shown in the schematic below, where S denotes a binding site, while a and b describe the relative affinities (a>0, b>0).



The fraction of any specific state (denoted by […]; for example, [S_2] denotes the fraction of the site bound by a dimer) can be obtained by solving equations:

$$([\boldsymbol{S\_2}]+[\boldsymbol{S\_4}]+[\boldsymbol{S\_6}])/\boldsymbol{\tau}-[\boldsymbol{S\_0}]\,\boldsymbol{n}/\boldsymbol{\tau}=\mathbf{0}$$

$$[\boldsymbol{S\_0}]\,\boldsymbol{n}/\boldsymbol{\tau}+[\boldsymbol{S\_4}]/(\boldsymbol{a\tau})+[\boldsymbol{S\_6}]/(\boldsymbol{b\tau})-[\boldsymbol{S\_2}]/\boldsymbol{\tau}-[\boldsymbol{S\_2}]\,\boldsymbol{n}/(\boldsymbol{a\tau})=\mathbf{0}$$

$$[\boldsymbol{S\_2}]\,\boldsymbol{n}/(\boldsymbol{a\tau})+[\boldsymbol{S\_6}]/(\mathbf{2}\boldsymbol{b\tau})-[\boldsymbol{S\_4}](\mathbf{1}/(\boldsymbol{a\tau})+\boldsymbol{n}/(\boldsymbol{b\tau})+\mathbf{1}/\boldsymbol{\tau})=\mathbf{0}$$

$$[\boldsymbol{S\_4}]\,\boldsymbol{n}/(\boldsymbol{b\tau})-[\boldsymbol{S\_6}](\mathbf{1}/\boldsymbol{\tau}+\mathbf{1}/(\boldsymbol{b\tau})+\mathbf{1}/(\mathbf{2}\boldsymbol{b\tau}))=\mathbf{0}$$

The results are:

$$[S\_0]=\frac{1}{1+n}$$

$$[S\_2]=\frac{n\left(3b+3ab+2b^2+2ab^2+2an+2abn\right)}{(1+n)\left(3b+3ab+2b^2+2ab^2+2an+3bn+2abn+2b^2n+2bn^2\right)}$$

$$[S\_4]=\frac{\left(3b+2b^2\right)n^2}{(1+n)\left(3b+3ab+2b^2+2ab^2+2an+3bn+2abn+2b^2n+2bn^2\right)}$$

$$[S\_6]=\frac{2bn^3}{(1+n)(3b+3ab+2b^2+2ab^2+2an+3bn+2abn+2b^2n+2bn^2)}$$

Clearly, changing a and b only affects the values of slopes rather than the monotonicity of these functions. Thus, a more detailed description with a and b restrained in proper ranges does not significantly affect the conclusions by the simplified description in Table S2.

S4. Statistical analysis of transcription initiation under different conditions

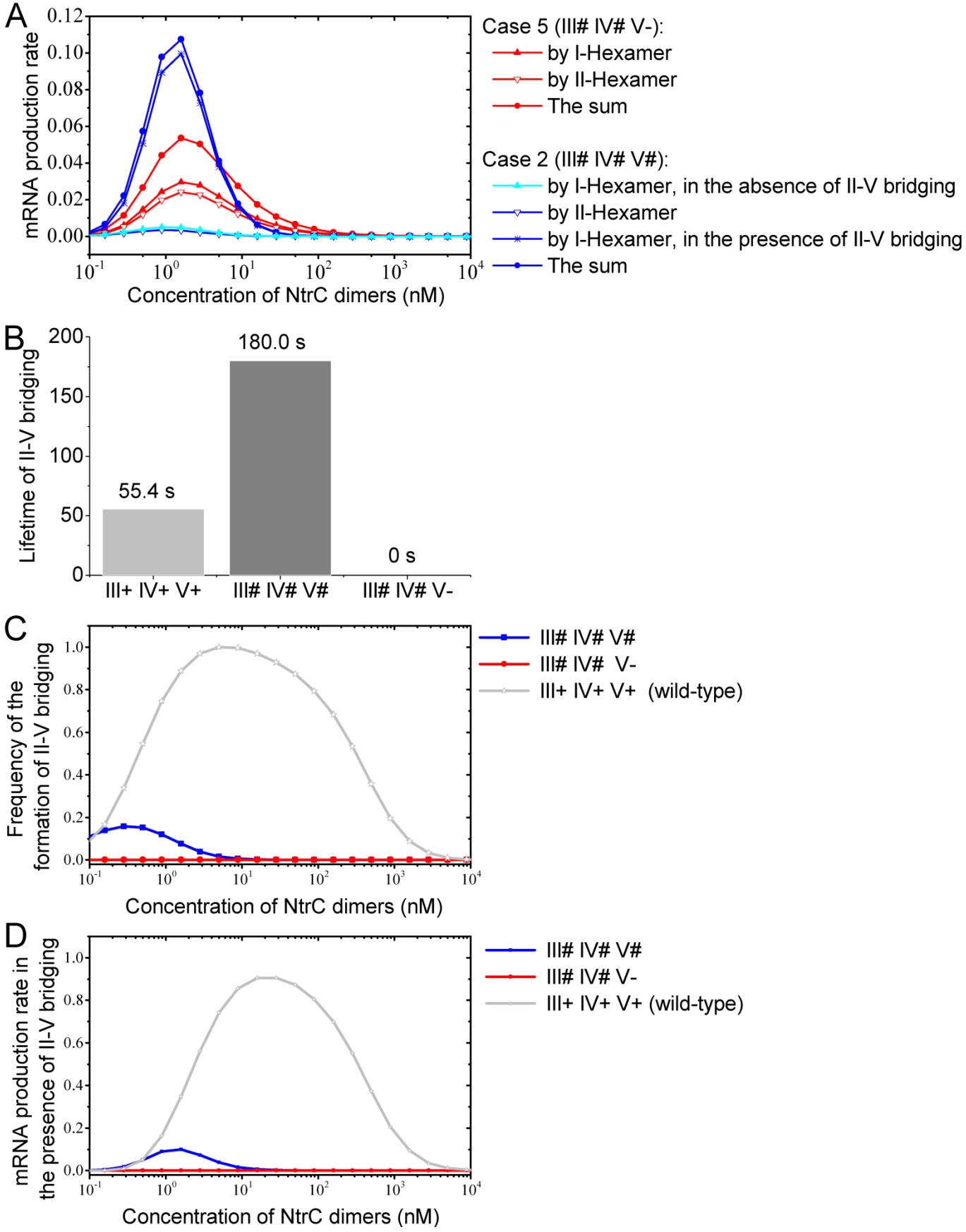


Figure S2. Statistical analysis of transcription initiation facilitated by the II-V bridging. (*A*) The

average rates of transcription initiations which are stimulated by enhancer II-bound hexamers, by enhancer I-bound hexamers in the presence or absence of II-V bridging, respectively. The data are normalized by setting the maximal transcription rate on the wild-type promoter to 1.0. (***B***) The average lifetime of II-V bridging in Case 2 (III# IV# V#) is much longer than that in the wild-type case, whereas no II-V bridging occurs in Case 5 (III# IV# V-). The lifetime is calculated based on $T = 1/(1/\tau_{II} + 2/\tau_o + 1/\tau_V)$, where τ_{II} and τ_V separately denote the average duration of a DBD bound to sites II and V, and τ_o denotes that of a dimer in a hexamer, equaling 720 s. In Case 2, τ_{II} and τ_V equal 720 s and 720 s; in the wild-type case, τ_{II} and τ_V equal 720 s and 72 s, respectively. (***C***) The times of forming the II-V bridging in a long time window are counted and are normalized by setting the maximum in the wild-type case to 1.0. (***D***) The rate of transcription initiation stimulated by enhancer I-bound hexamers in the presence of II-V bridging in Case 2 is much smaller than that in the wild-type case. The data are normalized by setting the maximum of total rate in the wild-type case to 1.0.

S5. Robustness of the stochastic model

To test the robustness of the model, we added Gaussian white noise to each rate constant of the 70 reactions. That is,

$$C_i' = C_i + \delta_i,$$

where C_i is the rate constant of reaction i $(i = 1, 2, \cdots, 70)$, and δ_i is a random number from the Gaussian distribution $N(0, 0.2\,C_i)$. The value of δ_i is independently assigned each time to update the number of reactants. The noise leads to a decrease in transcription rate for both the wild-type and mutated *glnAp2* promoters (III+ IV+ V+, III- IV- V-, and III# IV# V#). However, the *R*-*C* relation almost keeps unchanged (Figure S3A). That is, the basic characteristics of transcriptional behavior do not vary much, confirming that the model is sufficiently robust. Additionally, when white noise is only added to the rate constants of the holoenzyme interacting with DNA (i.e., those of reactions 61-74 and 67-70), the resulting *R*-*C* curves are shown in Figure S3B. The comparison of two panels in Figure S3 reveals that the changes in kinetics of the holoenzyme-DNA interactions largely contribute to the drop in transcription rate induced by noise. These changes prolong the time period from the holoenzyme binding to DNA to transcription initiation. In Case 2 (III# IV# V#) where the II-V bridging seldom forms but is rather stable once formed, however, the changes in the stability of II-V bridging also play a role (cf. Figure S2).

We also explored how the transcriptional behavior is affected by changes in those rate constants that are crucial for transcription initiation. They also turn out to be very robust (Figures 5 and S4-S8). In plotting Figure 5A, the default half-life of the II-V bridging is 55 s (τ_e =720 s; τ_l = 72 s; τ_o = 720 s). For the other values of half-life: 1.1 s (τ_e = 12 s; τ_l =1.2 s; τ_o = 720 s), 11 s (τ_e = 120 s; τ_l =12 s; τ_o = 720 s), 30 s (τ_e =360 s; τ_l =36 s; τ_o = 720 s), 96 s (τ_e =1440 s; τ_l =144 s; τ_o = 720 s), 232 s (τ_e =7200 s; τ_l =720 s; τ_o = 720 s). In plotting Figure 5B, only τ_o is varied given the half-life of a hexamer equals $\tau_o/3$. All the data in Figures S4-S8 are normalized by setting the maximal transcription rate in the wild-type case with default parameter values to 1.0.

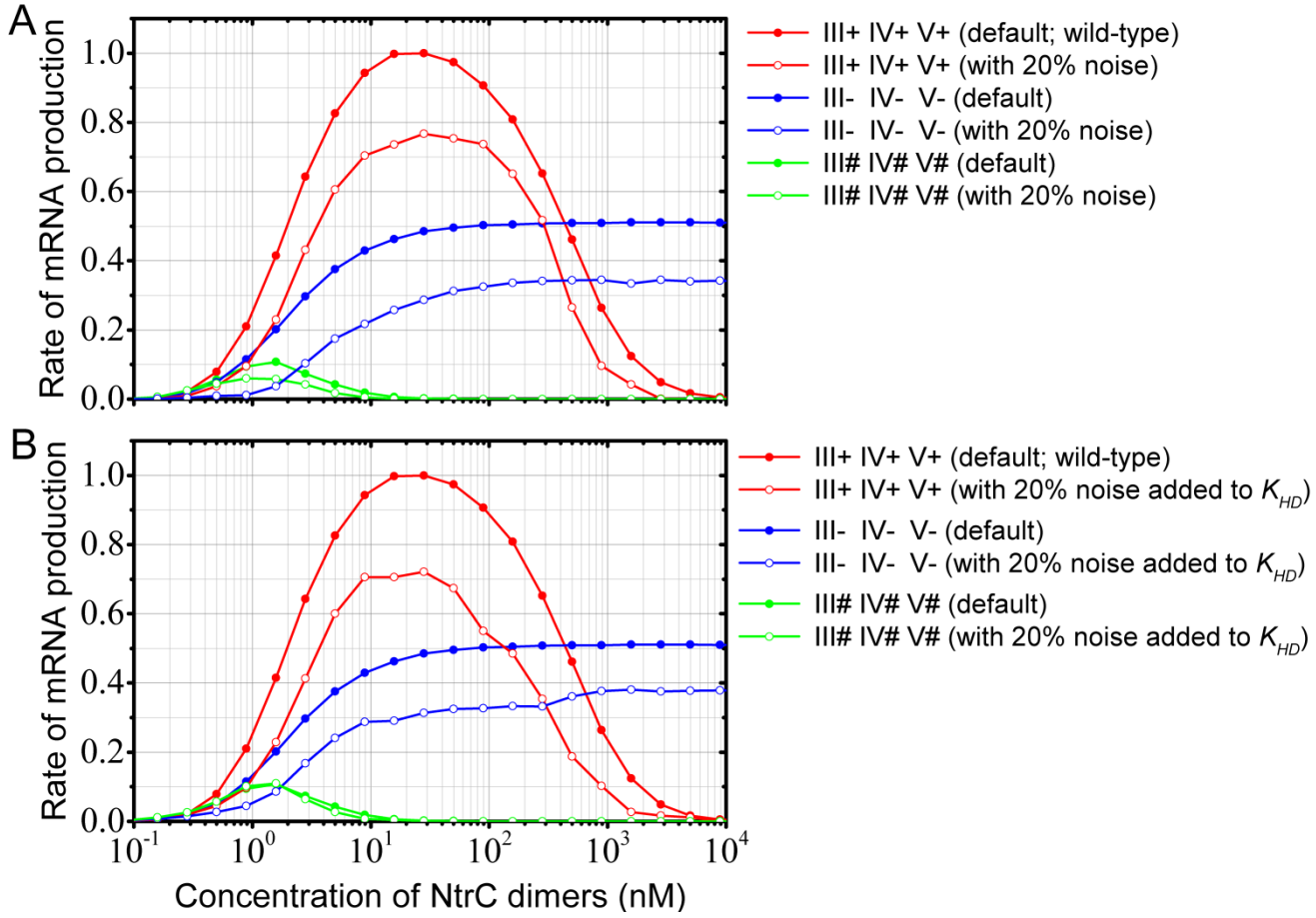


Figure S3. Robustness of the model. Addition of Gaussian white noise to rate constants of biochemical reactions leads to a decrease in transcription rate, without changing the relative dependence of R on C. In plotting (*B*), Gaussian white noise is only added to the rate constants of the holoenzyme interacting with DNA, and these constants are collectively referred to K_{HD}. Notably, the decrease is mainly due to the change in kinetics of the holoenzyme interacting with DNA. In the case of III# IV# V#, however, the decrease is largely due to a change in the stability of II-V bridging, which seldom forms but is rather stable once formed.

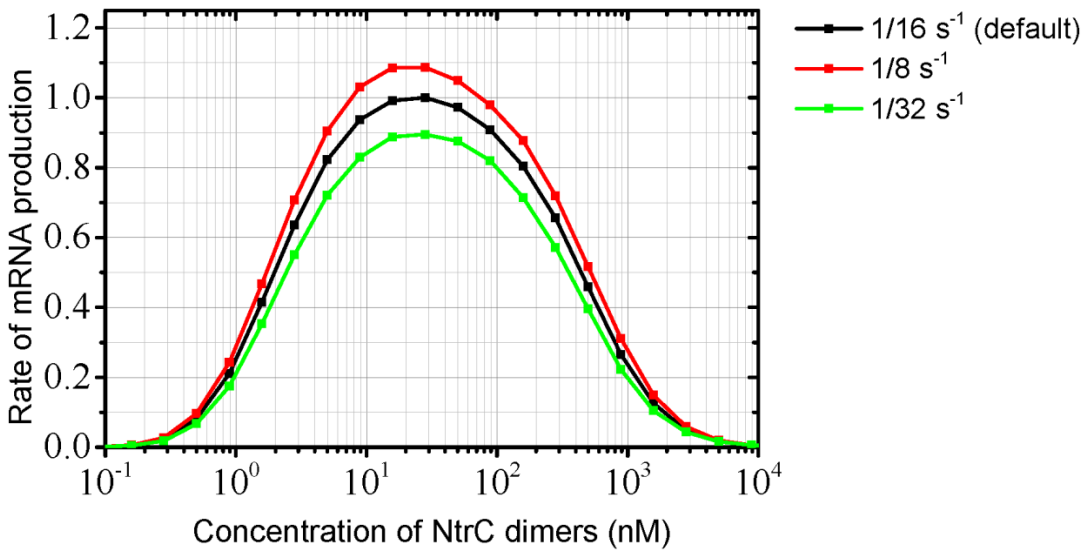


Figure S4. Halving or doubling the rate constant for forming the II-V bridging, $1/\tau_{25}$, leads to slight changes in the rate of mRNA production.

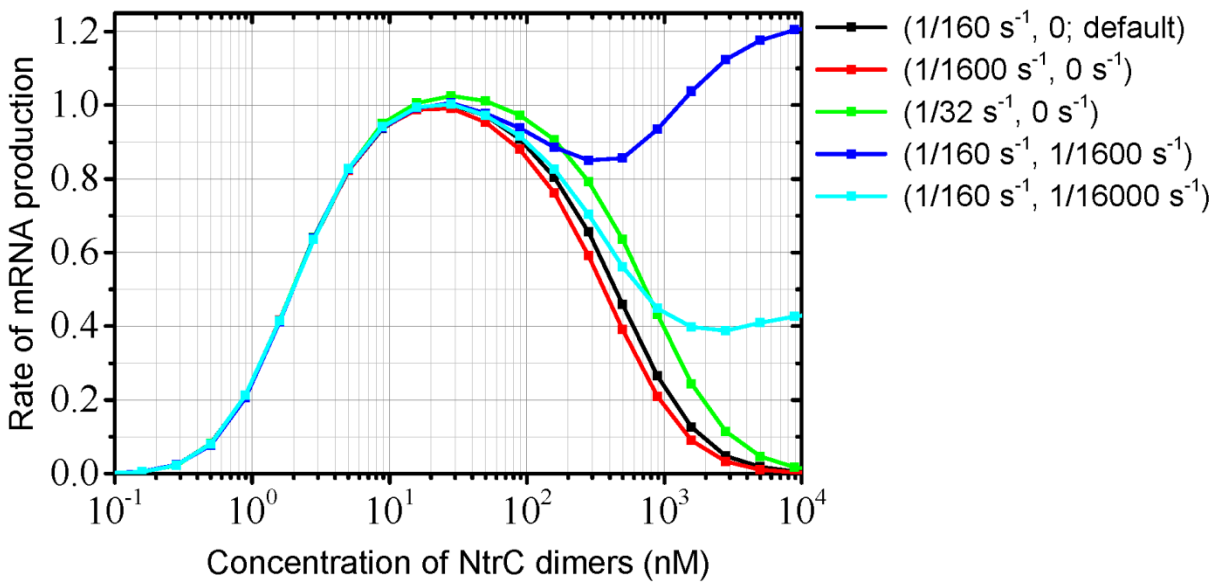


Figure S5. Effect of sites III and IV on transcription inhibition by hindering the formation of II-V bridging. The rates of reactions 46-52, which describe the formation of II-V bridging, depend on whether sites III and IV are occupied: the default rate constants are separately 1/16 s^{-1}, 1/160 s^{-1}, and 0 s^{-1} when neither, either, or both of sites III and IV are occupied. The two numbers in each bracket separately denote the rate constant when either site III or IV is occupied, or both are occupied. These data show that if both the two sites are occupied, the formation of II-V bridging should be substantially hindered; otherwise, the transcriptional behavior is no longer consistent with experimental observations.

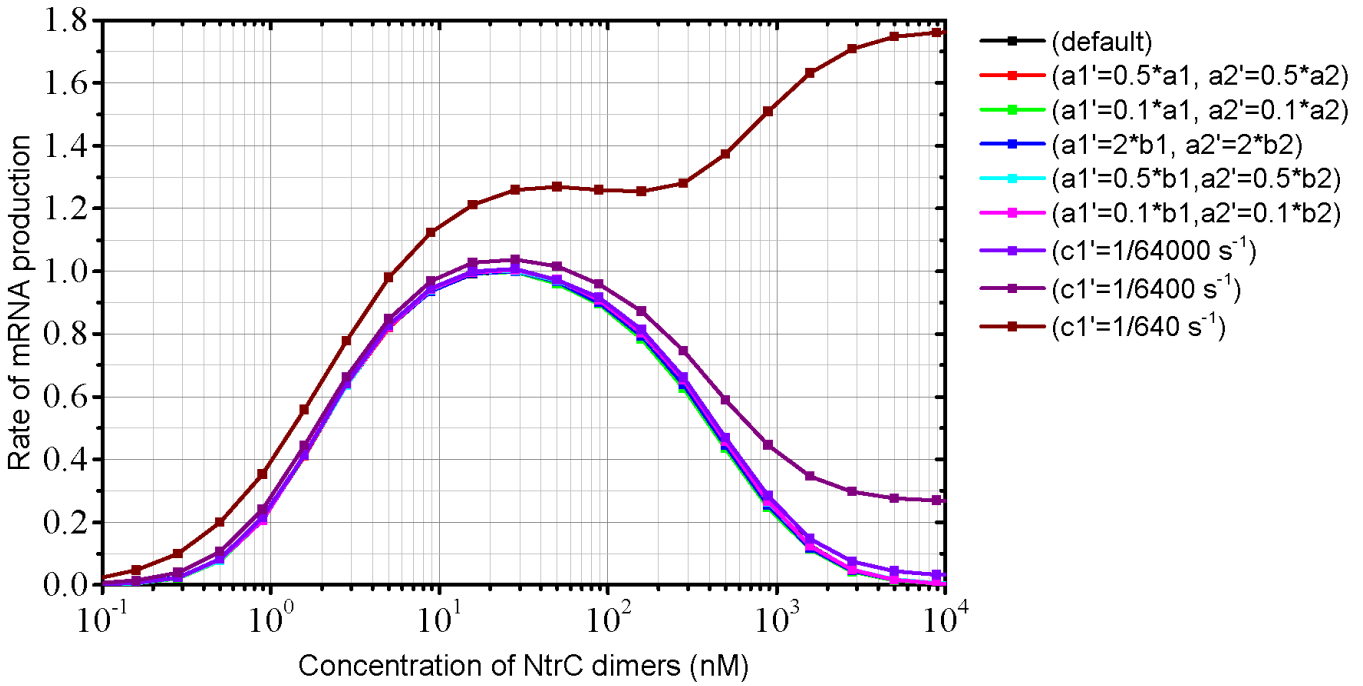


Figure S6. Effect of three low-affinity sites on repressing transcription by preventing hexamers at enhancers I and II from contacting the holoenzyme. If the three sites are bound by NtrC, the DNA looping is blocked and the enhancer I or II-bound hexamers are inhibited from contacting the holoenzyme. When one, two, or all of the three sites are occupied, the default rate constant of reaction 65 is separately *a*1, *b*1, and *c*1 (*a*1=1/128 s^{-1}, *b*1=1/640 s^{-1} and *c*1=*0* s^{-1}), and that of reaction 66 is separately *a*2, *b*2, and *c*2 (*a*2=1/160 s^{-1}, *b*2=1/640 s^{-1}, and *c*2=0 s^{-1}). The altered parameters are specifically denoted in the brackets. The transcriptional behaviors are very robust provided not all sites are occupied. If all the sites are occupied, enhancer I-bound hexamers should not contact the holoenzyme; otherwise, the transcriptional behavior is no longer consistent with experimental observations.

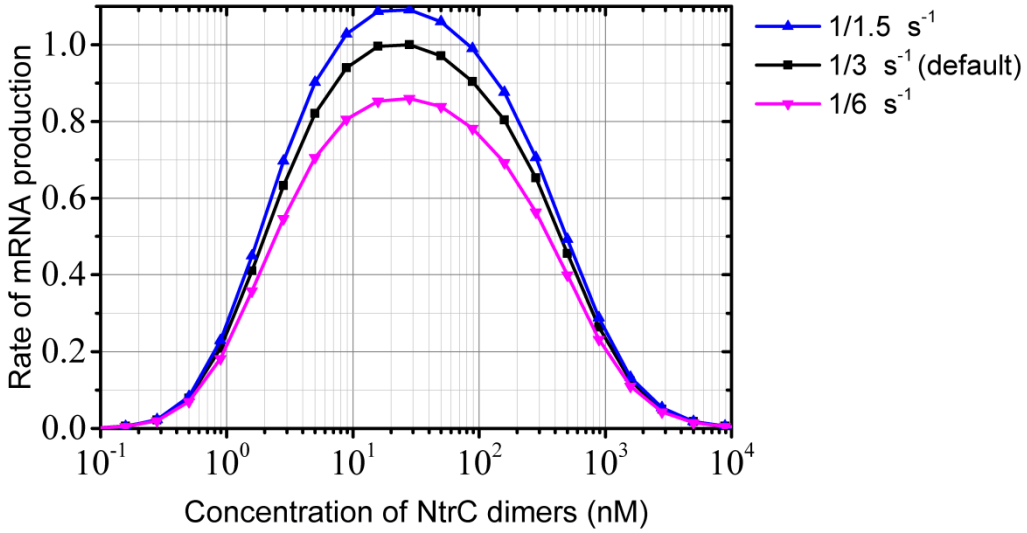


Figure S7. Slight changes in the transcriptional behavior are induced by halving or doubling the searching time for a hexamer at enhancer I to contact the holoenzyme given the II-V bridging. Given the II-V bridging, enhancer I-bound hexamers are just near the -12 region, where the transition to the open complex is catalyzed. The searching time is estimated to be 3 s by default. Halving or doubling this time leads to slight changes in the transcriptional behavior.

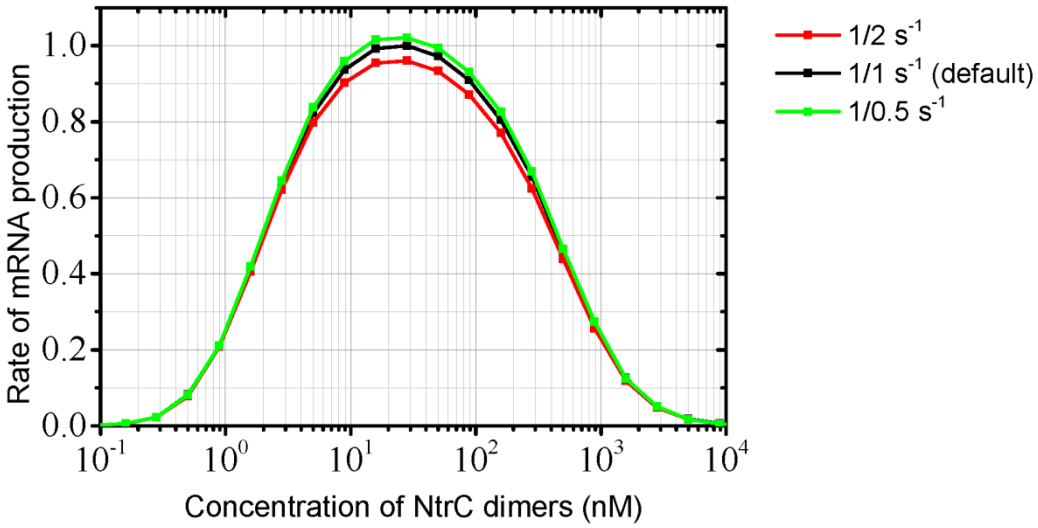


Figure S8. Halving or doubling the time of an enhancer-bound hexamer in contact with the holoenzyme leads to a slight change in the transcriptional behavior.

S6. Statistical analysis of the pathways leading to the II-V bridging

Pathway	% of pathways at individual concentrations of NtrC dimers						% of pathways for all concentrations						
	0.5 nM	2.8 nM	16 nM	89 nM	500 nM	2811 nM	0.5 nM	2.8 nM	16 nM	89 nM	500 nM	2811 nM	Sum
II_0 + V_6	3.0	4.7	3.9	1.5	0.6	0.0	0.4	1.3	1.0	0.3	0.1	0.0	3.1
II_0 + V_4	4.0	1.2	0.2	0.0	0.0	0.0	0.6	0.3	0.1	0.0	0.0	0.0	1.0
II_2 + V_2	3.6	0.8	0.2	0.0	0.0	0.0	0.5	0.2	0.1	0.0	0.0	0.0	0.8
II_2 + V_4	1.7	3.2	1.6	4.0	0.3	0.0	0.2	0.9	0.4	0.9	0.0	0.0	1.7
II_4 + V_0	56.2	18.9	4.3	0.8	0.6	0.0	8.1	5.1	1.2	0.2	0.1	0.0	14.6
II_4 + V_2	1.9	4.1	1.6	0.6	0.0	0.0	0.3	1.1	0.4	0.1	0.0	0.0	2.0
II_6 + V_0	29.5	67.0	88.1	96.6	98.6	100	4.3	17.9	23.7	20.6	9.5	0.9	76.9

Table S3. Statistical analysis of pathways leading to the II-V bridging. At rather low concentrations (around 0.5 nM), the formation of II-V bridging is triggered mainly by enhancer II-bound NtrC tetramers contacting site V. Beyond that, the formation is predominantly induced by enhancer II-bound NtrC hexamers contacting site V.

S7. Additional predictions

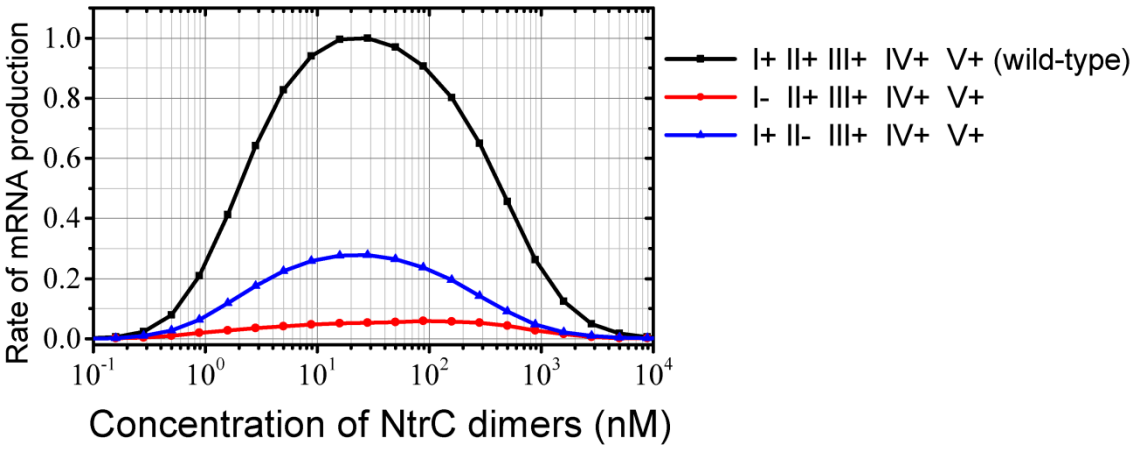


Figure S9. Predicted transcriptional activities from two mutated promoters. When either enhancer I or enhancer II is mutated to a sequence that does not bind any protein, the transcription rate drops substantially over a wide range of NtrC concentration in comparison with the wild-type case. Additionally, the transcription rate is much lower when enhancer I is mutated.

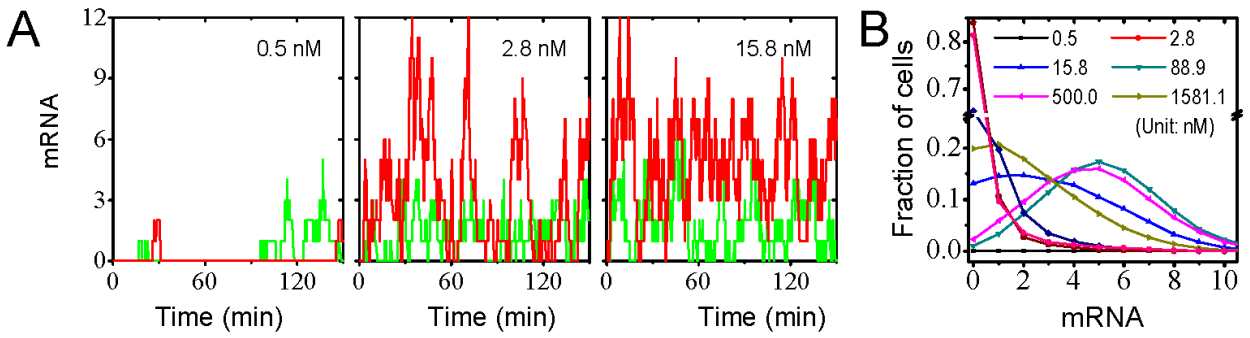


Figure S10. Temporal evolution of mRNA synthesis and distribution of the number of mRNAs. (***A***) Temporal evolution of the number of mRNA transcripts from the wild-type (red) or mutated (III- IV- V-, green) promoter. The concentration of NtrC dimers is constantly kept at 0.5, 2.8 or 15.8 nM. The average lifetime of mRNA is assumed to be 150 s. (***B***) Distributions of the number of mRNA transcripts from the wild-type promoter over a cell population at various concentrations of NtrC dimers. The number is counted in the steady state long after the transient.

S8. References

1 De Carlo, S., et al. (2006) The structural basis for regulated assembly and function of the transcriptional activator NtrC. *Genes Dev.* **20**, 1485-1495.

2 Haykinson, M.J. and Johnson, R.C. (1993) DNA looping and the helical repeat in vitro and in vivo: effect of HU protein and enhancer location on Hin invertasome assembly. *EMBO J.* **12**, 2503-2512.

3 Lee, D.H. and Schleif, R.F. (1989) In vivo DNA loops in araCBAD: size limits and helical repeat. *Proc. Natl. Acad. Sci. U.S.A.* **86**, 476-480.

4 Porter, S.C., North, A.K., Wedel, A.B., and Kustu, S. (1993) Oligomerization of NTRC at the *glnA* enhancer is required for transcriptional activation. *Genes Dev.* **7**, 2258-2273.

5 Weiss, V., Claverie-Martin, F., and Magasanik, B. (1992) Phosphorylation of nitrogen regulator I of Escherichia coli induces strong cooperative binding to DNA essential for activation of transcription. *Proc. Natl. Acad. Sci. U.S.A.* **89**, 5088-5092.

6 Sevenich, F.W., Langowski, J., Rippe, K., and Weiss, V. (1998) DNA binding and oligomerization of NtrC studied by fluorescence anisotropy and fluorescence correlation spectroscopy. *Nucleic Acids Res.* **26**, 1373-1381.

7 Wiggins, P.A., et al. (2006) High flexibility of DNA on short length scales probed by atomic force microscopy. *Nat. Nanotechnol.* **1**, 137-141.

8 Zhang, J., Lin, M., Chen, R., Wang, W., and Liang, J. (2008) Discrete state model and accurate estimation of loop entropy of RNA secondary structures. *J. Chem. Phys.* **128**, 125107.

9 Lilja, A.E., Jenssen, J.R., and Kahn, J.D. (2004) Geometric and dynamic requirements for DNA looping, wrapping and unwrapping in the activation of *E. coli glnAp2* transcription by NtrC. *J. Mol. Biol.* **342**, 467-478.

10 Wales, D.J. and Salamon, P. (2014) Observation time scale, free-energy landscapes, and

molecular symmetry. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 617-622.

11 Jacobson, H. and Stockmayer, W.H. (1950) Intramolecular reaction in polycondensations. I. The theory of linear systems. *J. Chem. Phys.***18**, 1600-1606.

12 Friedman, L.J. and Gelles, J. (2012) Mechanism of transcription initiation at an activator-dependent promoter defined by single-molecule observation. *Cell* **148**, 679-689.

13 Huo, Y.X., et al. (2006) Protein-induced DNA bending clarifies the architectural organization of the σ^{54}-dependent *glnAp2* promoter. *Mol. Microbiol.* **59**, 168-180.

14 Zhang, Z., Revyakin, A., Grimm, J.B., Lavis, L.D., and Tjian, R. (2014) Single-molecule tracking of the transcription cycle by sub-second RNA detection. *Elife* **3**, e01775.

15 Atkinson, M.R., Pattaramanon, N., and Ninfa, A.J. (2002) Governor of the *glnAp2* promoter of Escherichia coli. *Mol. Microbiol.* **46**, 1247-1257.

16 Gillespie, D.T. (1976) A general method for numerically simulating stochastic time evolution of coupled chemical reactions. *J. Comput. Phys.* **22**, 403-434.

17 Gillespie, D.T. (1977) Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.* **81**, 2340-2361.

18 Wang, Y., Liu, F., and Wang, W. (2012) Dynamic mechanism for the transcription apparatus orchestrating reliable responses to activators. *Sci. Rep.* **2**, 422.