

SCIENTIFIC REPORTS



OPEN

The free-energy cost of interaction between DNA loops

Lifang Huang^{1,2}, Peijiang Liu², Zhanjiang Yuan³, Tianshou Zhou³ & Jianshe Yu¹

From the viewpoint of thermodynamics, the formation of DNA loops and the interaction between them, which are all non-equilibrium processes, result in the change of free energy, affecting gene expression and further cell-to-cell variability as observed experimentally. However, how these processes dissipate free energy remains largely unclear. Here, by analyzing a mechanic model that maps three fundamental topologies of two interacting DNA loops into a 4-state model of gene transcription, we first show that a longer DNA loop needs more mean free energy consumption. Then, independent of the type of interacting two DNA loops (nested, side-by-side or alternating), the promotion between them always consumes less mean free energy whereas the suppression dissipates more mean free energy. More interestingly, we find that in contrast to the mechanism of direct looping between promoter and enhancer, the facilitated-tracking mechanism dissipates less mean free energy but enhances the mean mRNA expression, justifying the facilitated-tracking hypothesis, a long-standing debate in biology. Based on minimal energy principle, we thus speculate that organisms would utilize the mechanisms of loop-loop promotion and facilitated tracking to survive in complex environments. Our studies provide insights into the understanding of gene expression regulation mechanism from the view of energy consumption.

Transcription is a pivotal but also the most complex step during gene expression. Transcription of a gene is regulated often by promoter-proximal DNA elements and distal DNA elements that together determine the expression pattern of this gene. Enhancers activate promoters by directly contacting binding sites for transcription factors and chromatin-modifying enzymes via DNA looping¹⁻⁵. Chromatin elements may form DNA loops of spatial structure in different manners, e.g., promoter-tethering elements in *Drosophila* that allow activation by specific enhancers over long distances are proposed to form DNA loops between sequences near the enhancer and the promoter^{6,7}. Also for example, in bacteriophage λ , the CI protein forms a 2.3-kb DNA loop that brings a distal stimulatory site close to RNA polymerase at the PRM promoter⁸. Li, *et al.*⁹ thought that one major factor determining the formation of a DNA loop is the flexibility of chromatin, which may be modulated by histone acetylation and other modifications¹⁰. The specific interactions between chromatin elements can either assist enhancer-promoter looping by bringing the enhancer and the promoter closer together, or are thought to interfere with enhancer-promoter looping by placing them in separate loop domains¹¹. In a more general sense, chromatin elements are linked in a highly complex network of DNA-looping interactions, where enhancers often (e.g., in eukaryotes) directly contact promoters over large genomic distances to regulate gene expression. Characterizing the principles underlying these distal enhancer-promoter contacts is fundamentally important for a full understanding of gene expression.

Revealing the mechanisms of how transcriptional enhancers control over 20,000 protein-coding genes to maintain cell-specific gene expression programs is a fundamental challenge in biology¹². Studies¹²⁻¹⁷ suggested complex promoter structures formed by the interaction of DNA sites, especially by enhancer-gene interactions. These interactions are in essential biochemical, leading to stochastic transcription and further cell-to-cell variability^{18,19}. In recent years, the study of chromosomal loop structures and loop-loop interactions has been receiving increasing attention. Priest, *et al.*, used two well-characterized DNA-looping proteins¹¹: Lac repressor and phage λ CI, to measure interactions between pairs of long DNA loops in *E. coli* cells in three possible topological patterns of two pairs of interacting sites on DNA, namely, side-by-side loops, nested loops, and alternating loops. They found that two DNA loops in the side-by-side structure do not affect each other; those in the nested

¹Research Centre of Applied Mathematics, Guangzhou University, Guangzhou, 510006, P.R. China. ²School of Statistics and Mathematics, Guangdong University of Finance & Economics, Guangzhou, 510275, P.R. China.

³Guangdong Province Key Laboratory of Computational Science, School of Mathematics and Computational Science, Sun Yat-Sen University, Guangzhou, 510275, P.R. China. Correspondence and requests for materials should be addressed to T.Z. (email: mcszhtsh@mail.sysu.edu.cn) or J.Y. (email: jsyu@gzhu.edu.cn)

structure assist each other's formation consistent with the distance-shortening effect; and those in the alternating structure, where one looping element is placed within the other DNA loop, inhibit each other's formation, thus providing clear support for the loop domain model for insulation. In addition, they argued that the combination of loop assistance and loop interference can provide strong specificity in long-range interactions. Another similar but important work is that Savitskaya, *et al.*, experimentally observed¹³ that when a pair of repressor (Su and Hw) found in the gypsy retrotransposon and their binding sites are in between the enhancer and the promoter, the gene expression level is not decreased but increased. For this counter-intuitive phenomenon, Mirny, *et al.*, conjectured that the Su and Hw pair shortens the distance between the enhancer and the promoter, leading to the rise of the expression level¹⁴. Despite these studies, it is still unclear to what extent this mechanism contributes to specific enhancer-gene interactions¹² and which factors influence as well as how they affect the interaction between DNA loops. In this paper, we will address these questions from a viewpoint of energy dissipation. Our study is motivated by two important questions: The one is on the communication form between enhancer and promoter, namely, which of the direct looping and the facilitated-tracking looping has advantages^{9,15,20}. The other is how DNA loops interact with one another, whether one DNA loop modulates the expression of another loop by changing the transcriptional rate^{21–23}. Experiment data on the measured dependency of noise upon average expression can well fit the variations in transcriptional rates across chromosome²¹, but cannot specify the interaction between DNA loops and its effect on gene expression. This article tries to understand the interaction between DNA loops from the viewpoint of the minimal energy principle. For clarity, we will restrict our study to two representative interacting DNA loops (the one formed by silence genes, e.g., Su and Hw¹³, or Anchor and CTCF¹² and the other by expressing genes, e.g., an enhancer and a promoter) that altogether may form the three fundamental structures as mentioned above.

As a fact, maximizing the energy available to cells for biosyntheses, growth, and division is essential for cell survival²⁴, whereas the operation mechanisms of biological systems usually follow the minimal energy principle with^{4–28}. For example, Wong, *et al.* showed that intron retention can regulate the activity of many genes, and this requires only a small amount of energy²⁹. In addition, Wang, *et al.* showed that the path of differentiation in stem cell is obtained along the direction of minimum energy³⁰. These works imply that using the minimal energy principle to infer biological mechanisms would be reasonable. In our case, the DNA loop formation is essentially a non-equilibrium process from the viewpoint of statistical thermodynamics, thus necessarily consuming energy. In fact, storage and retrieval of the genetic information in a cell are a dynamic process that requires the DNA to undergo structural rearrangements essential for transcriptional regulation in prokaryotic and eukaryotic cells³¹, DNA looping and the interaction between DNA loops are two prominent examples of such a structural deformation. This deformation belonging to energy-dependent chromatin remodeling necessarily dissipates energy. Interestingly, Chen, *et al.*³¹ showed that both loop association and loop dissociation at the DNA-repressor junctions depend on the elastic deformation of the DNA and protein, and that both looping and unlooping rates approximately scale with a looping factor that can well reflect the system's deformation free energy. They found that the loop breakdown process at the DNA-protein interface is sensitive to the whole loop's deformation, and that both looping and unlooping kinetics exhibit rather simple forms of scaling with the looping free energy. In addition, Coulon *et al.*³² studied a general model of stochastic gene expression that can represent arbitrary prokaryotic or eukaryotic promoters, and revealed the potential activity of any promoter and its influence on gene expression from the perspective of energetic cost (e.g., they showed that the regime where molecular interactions underlying promoter dynamics have typical characteristics such as high mobility, functional redundancy and many alternate states/pathways results in direct and indirect energetic cost). The minimal energy principle would provide a good angle of view for understanding the regulatory mechanism of the interaction of DNA loops, but because of the complexity of biological regulations, energy calculation is nontrivial and needs to develop new methods.

We notice that the study of Chen, *et al.*³¹ and other similar works^{33–35} studied free energy dissipation but considered only the case of single DNA loops. Thus far, how the interaction between DNA loops consumes free energy and the relationship of the gene expression level with energy dissipation is not clear. In addition, classical thermodynamic models³⁶ (used in the above mentioned works), which are based on the assumption that the level of gene expression is proportional to the equilibrium probability that RNA polymerase (RNAP) is bound to the promoter of interest, provide only a framework for understanding energy dissipation in gene expression, and in general cannot specify which important processes (e.g., the formation of DNA loops and their interaction) that can efficiently regulate gene expression levels contribute to energy consumption in the full system.

Here, we develop a new but intuitive computational approach to analyze the free energy cost of the interaction between chromatin loops. Specifically, we first map complex topological patterns for the interactions among DNA loops into a multistate model of stochastic transcription, where transition rates between promoter states are functions of the loop lengths (along DNA lines). Then, we propose a principle of free energy consumption (defined as the entropy production rate $\dot{W} = dW/dt$ in this paper, also called the dissipation rate of free energy), which decomposes this dissipation rate into two parts: the one for promoter dynamics (denoted by \dot{W}_p) and the other for promoter-mediated transcription dynamics (\dot{W}_y). The former \dot{W}_p is easily calculated using the standard statistical thermodynamics method^{37–39}, while the latter \dot{W}_y can be simply estimated from Fokker-Planck equations for the mRNA-concentration distributions with each derived by considering that the gene is at a single state (see the below Methods for details). Our energy dissipation model also can conveniently incorporate other factors such as the communication form between regulatory elements^{12,13,16}. Importantly, apart from showing that a longer DNA loop needs more mean free energy consumption (defined as the ratio of the entropy production rate over the mean mRNA), we find that the promotion of two loops always reduces the mean free energy consumption whereas the suppression has the opposite effect, both independent of the topology of two DNA loops; in contrast to the mechanism of direct looping between regulatory elements, the facilitated-tracking mechanism dissipates

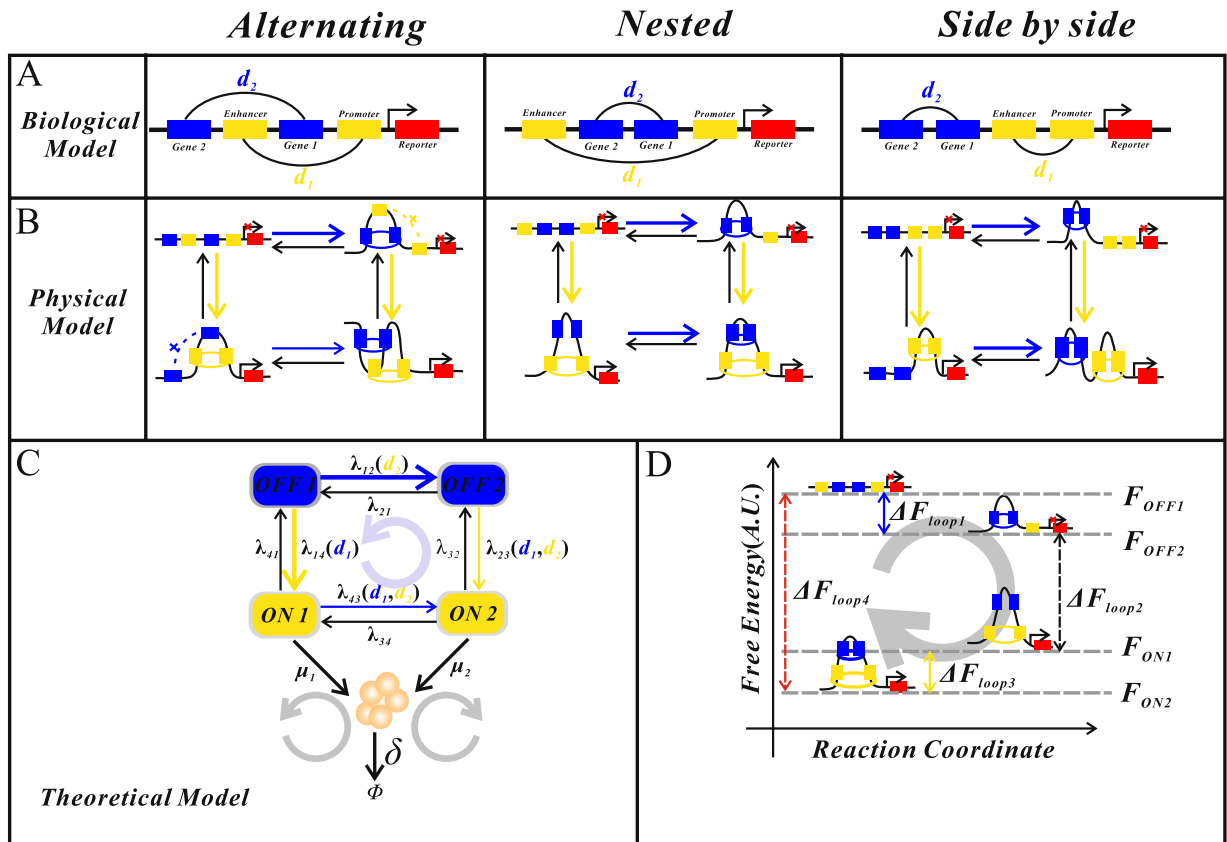


Figure 1. Schematic diagram for two interacting DNA loops. **(A)** Three fundamental biological structures, where two pairs: Gene1/Gene2 (yellow dock), enhancer/promoter (blue dock), may form two distinct topologies. **(B)** Physical structures for respective DNA–looping interactions in **(A)**, which consider the possibility of looping. **(C)** Theoretical model by mapping the physical models into a 4-state model of gene expression, where transition rates between active and inactive states may depend the loop lengths along the DNA lines (called looping rates), and the grey loop with arrow represents free energy flow. **(D)** Free energy difference between unlooped and looped states.

less mean free energy and can enhance the mean mRNA expression, justifying the facilitated-tracking hypothesis in biology.

Materials and Methods

Hypotheses based on experimental evidence. In order to reveal the essential mechanism of how interacting DNA loops consumes free energy, here we consider only the interaction between two loops: the one formed by a pair of insulators: Su/Hw¹³ or Anchor/CTCF¹², and the other formed by enhancer and promoter. For convenience, the former loop is denoted as the blue loop whereas the latter loop by the yellow loop, referring to Fig. 1A. According to ref.¹¹, there are three possible connection topologies between these two pairs of DNA loops: cross-type structure (due to alternating loops), inline-type structure (due to nested loops), and independence-type structure (due to side-by-side loops).

Denote by d_1 the length of the yellow loop along the DNA line, and by d_2 the length of the blue loop also along the DNA line. Experimental evidence supports that alternating loops give loop interference, nested loops give loop assistance, and side-by-side loops do not interact^{11,13}. We assume that the gene is expressed only after the yellow loop is formed. Note that in theory, a pair of DNA regulatory elements may form a loop but also may not form any loop, i.e., there are two possibilities. Thus, there are in total four possibilities for each of the three topologies, referring to Fig. 1B. To help the readers understand this schematic figure, we state additional details: (1) If the yellow loop is formed but the blue loop is not formed, then the gene is expressed. Moreover, gene expression can be enhanced; (2) If both the yellow loop and the blue loop are formed, then the gene is also expressed. However, the transcriptional rate may be different from that in the former case since experiment data²¹ suggested that the formation of one DNA loop modulate the transcriptional rate of another DNA loop. In addition, the expression effect may also be different between cross-type and inline-type structures. Specifically, for the former, the formation of the blue loop represses the effect that the yellow loop enhances gene expression, whereas for the latter, the case is just opposite; (3) If neither the yellow loop nor the blue loop is formed, then the gene is not expressed; (4) If the yellow loop is not formed but the blue loop is formed, then the gene is not expressed either.

Next, we map the three physical models in Fig. 1B into a common multistate model of gene expression, referring to Fig. 1C. With this mapping, a complex question of how two interacting DNA loops affects gene expression

is transformed to a simple one of how a 4-state model of stochastic transcription is solved. Note that after mapping, the looping rates as functions of loop lengths currently become transition rates between promoter activity states. In addition, once two DNA loops are formed, any one of them can impact the length of the other, often in a nonlinear manner. Moreover, this impact can lead to changes in transition rates and further in the mRNA level. Also note that the ON1 and ON2 states indicated in Fig. 1C mean that the enhancer and promoter pair forms a loop (corresponding to the yellow loop) whereas the other pair of elements may form a loop (corresponding to the blue loop) but also may not form any loop. The transcriptional rates in ON1 and ON2 states may be different due to the interaction between two DNA loops. In contrast, at OFF1 and OFF2 states, the enhancer and the promoter pair does not form a loop whereas the other pair of regulatory elements may form a loop but also may not form any loop.

According to the above mapping relationships between physical and theoretical models, we know that the rates of loop dissociation and association, λ_{12} , λ_{21} , λ_{23} , λ_{32} , λ_{43} , λ_{34} , λ_{14} , λ_{41} (unit: bp/second), are transformed to transition rates between active and inactive states shown in Fig. 1C. The former rates depend on the lengths or distances of the two loops (along the DNA lines), so do the latter rates. In the case of single DNA loops, previous works gave exponential formulae for the relationship between the looping rate (unit: bp/second) and the loop length (unit: bp)^{19,36}. In our case, these formulae read

$$\begin{aligned}\lambda_{14} &= k_{loop}^{(1)} = \beta \exp\left(-\frac{u}{d_1} - v \log(d_1) + wd_1 + z\right) \\ \lambda_{12} &= k_{loop}^{(2)} = \beta \exp\left(-\frac{u}{d_2} - v \log(d_2) + wd_2 + z\right)\end{aligned}\quad (1)$$

where $k_{loop}^{(i)}$ represent the DNA looping rates, d_i are the loop lengths along the DNA lines, and u, v, w, z are parameters, the values of which can be obtained by fitting experimental data^{19,36}, e.g., $u = 140.6$, $v = 2.52$, $w = 0.0014$, and $z = 19.9$. Parameter β is a normalized constant for which we set $\beta = 1/1000$ throughout this article.

In general, each of two transition rates, λ_{23} and λ_{43} is a function of two distances, d_1 and d_2 . However, the existing experimental data support only the quantitative relationship between the DNA looping rate and the length of the yellow loop¹⁴. Based on the above analysis and without loss of generality, we may set $\lambda_{23} = k_1 \lambda_{14}$, $\lambda_{43} = k_2 \lambda_{12}$, where parameters k_1 and k_2 are set, based on experimental data¹⁴ with small modifications, as

$$k_1 = \begin{cases} 0.73 + \frac{1}{d_1} < 1 & \text{Alternating loops} \\ 40 e^{-0.05d_1} + 1 > 1 & \text{Nested loops} \\ 1 & \text{Side-by-side loops} \end{cases}\quad (2a)$$

$$k_2 = \begin{cases} 0.73 + \frac{1}{d_2} < 1 & \text{Alternating loops} \\ 4 & \text{Nested loops} \\ 1 & \text{Side-by-side loops} \end{cases}\quad (2b)$$

Finally, we point out the following points: (1) DNA loops exist extensively, especially in ukaryotes^{11–14}; (2) there are three representative types of interactions between DNA loops (alternating, nested and side-by-side loops), each supported by experimental data¹¹; (3) there are two mainstream ways between DNA regulatory elements (direct looping and facilitated tracking looping), each supported also by experimental evidences^{9,15,20}; (4) dependence of looping rates on loop lengths is obtained by fitting experimental data^{19,36}; (5) our gene models that consider the interactions between DNA loops do not explicitly consider the regulatory roles of transcription factors (TFs), but if we let model parameters such as looping rates (or transition rates), and transcriptional rates change in their respective yet biologically reasonable ranges, then our models do not lost generality and in particular, they can capture the regulatory effects of TFs. This simplification, which has been adopted in many references^{40–42}, is here made for analysis convenience. Moreover, one will see that the results obtained in this paper are qualitatively unchanged, independent of the choice of parameter values, so the simplification is reasonable.

An approximate method for calculating mRNA probability distribution. One will see that energetic cost for the system under consideration has a closed relation with the probability distribution of the mRNA molecule number. In order to calculate this distribution, here we propose a simple and intuitive yet effective method, which is based on the isothermal decomposition of probability. Let x_1 , x_2 , x_3 and x_4 represent the DNA proportions (or fractions) at states OFF1, OFF2, ON2 and ON1, respectively; y represent the mRNA concentration. Denote by μ_1 and μ_2 the mRNA synthesis rates at ON1 and ON2 states respectively (unit: $\mu M/\text{sec}$); and by δ the mRNA degradation rate (unit: $\mu M/\text{sec}$). The deterministic equations for the full reaction system take the form

$$\begin{aligned}
\frac{dx_1}{dt} &= -(\lambda_{12} + \lambda_{14})x_1 + \lambda_{21}x_2 + \lambda_{41}x_4 \\
\frac{dx_2}{dt} &= -(\lambda_{21} + \lambda_{23})x_2 + \lambda_{12}x_1 + \lambda_{32}x_3 \\
\frac{dx_3}{dt} &= -(\lambda_{32} + \lambda_{34})x_3 + \lambda_{23}x_2 + \lambda_{43}x_4 \\
\frac{dx_4}{dt} &= -(\lambda_{43} + \lambda_{41})x_4 + \lambda_{34}x_3 + \lambda_{14}x_1 \\
\frac{dy}{dt} &= \mu_1x_4 + \mu_2x_3 - \delta y
\end{aligned} \tag{3}$$

where $x_1 + x_2 + x_3 + x_4 = 1$, Solving Eq. (3) at steady state, we obtain the following expressions

$$x_1^s = \frac{A}{E}, x_2^s = \frac{B}{E}, x_3^s = \frac{C}{E}, x_4^s = \frac{D}{E}, y^s = \frac{\mu_1 D + \mu_2 C}{\delta E} \tag{4}$$

with $A = \lambda_{21}\lambda_{32}(\lambda_{41} + \lambda_{43}) + \lambda_{34}\lambda_{41}(\lambda_{21} + \lambda_{23})$, $C = \lambda_{12}\lambda_{23}(\lambda_{41} + \lambda_{43}) + \lambda_{14}\lambda_{43}(\lambda_{21} + \lambda_{23})$, $B = \lambda_{12}\lambda_{32}(\lambda_{41} + \lambda_{43}) + \lambda_{12}\lambda_{34}\lambda_{41} + \lambda_{14}\lambda_{32}\lambda_{43}$, $D = \lambda_{14}\lambda_{21}\lambda_{32} + \lambda_{12}\lambda_{23}\lambda_{34} + \lambda_{14}\lambda_{34}(\lambda_{21} + \lambda_{23})$, $E = A + B + C + D$, which are all the implicit functions of DNA loop lengths along the underlying DNA lines.

In order to derive the mRNA probability distribution in an intuitive manner, we consider 4 extreme cases. The time scale of DNA looping is in general slower compared to that of transcription, so if the gene is only at OFF1 state, then the mRNA always degrades without production, implying that the mRNA concentration follows an exponential distribution. Specifically, if we denote by $P_1(y)$ the mRNA distribution in this case, then $P_1(y) = (A)/(E)\delta e^{-\delta y}$, where A/E is a weight. Similarly, the mRNA distribution only at OFF2 state, $P_2(y)$, is given by $P_2(y) = (B)/(E)\delta e^{-\delta y}$. If the gene is only at ON2 state, then the mRNA is both produced and degraded, implying that the mRNA concentration, denoted by $P_3(y)$, follows a Poisson distribution. From a mathematical view, Poisson distribution and normal distribution can be approximated to each other, implying that $P_3(y) = \frac{C}{E} \frac{1}{\sqrt{2\pi}\sigma_1} \exp\left[-\frac{(y - \Lambda_1)^2}{2\sigma_1^2}\right]$, where $\Lambda_1 = \frac{(\mu_1)}{\delta}$ and $\sigma_1^2 = \frac{\mu_1}{\delta}$, C/E is a weight. Similarly, the mRNA concentration only at ON1 state follows a normal distribution with a weight, i.e., $P_4(y) = \frac{D}{E} \frac{1}{\sqrt{2\pi}\sigma_2} \exp\left[-\frac{(y - \Lambda_2)^2}{2\sigma_2^2}\right]$, where $\Lambda_2 = \frac{(\mu_2)}{\delta}$ and $\sigma_2^2 = \frac{\mu_2}{\delta}$. Since the gene must be and is only at one of 4 states, the total protein distribution at steady state, denoted by $P(y)$, should be equal to the sum of the above 4 factorial distributions, that is, $P_1(y) = P_1(y) + P_2(y) + P_3(y) + P_4(y)$. Thus, we obtain the following analytical probability distributions of the mRNA concentration at steady state

$$\begin{aligned}
P(y) &= \frac{A}{E}\delta e^{-\delta y} + \frac{B}{E}\delta e^{-\delta y} + \frac{C}{E} \frac{1}{\sqrt{2\pi}\sigma_1} \exp\left[-\frac{(y - \Lambda_1)^2}{2\sigma_1^2}\right] \\
&\quad + \frac{D}{E} \frac{1}{\sqrt{2\pi}\sigma_2} \exp\left[-\frac{(y - \Lambda_2)^2}{2\sigma_2^2}\right]
\end{aligned} \tag{5}$$

This explicit expression is in good accordance with the one obtained by the Gillespie stochastic simulation algorithm^{43,44} (referring to Fig. 2A, where we have used the fact that the size of the probability distribution $P(x)$ at $x = i$ is equal to that of the area bounded by the corresponding probability density curve and the interval $[i - 1/2, i + 1/2]$), implying that the above approximation is effective. In other words, the total probability density is equal to the sum of the individual probability densities at discrete states.

Quantifying the free energy cost for the formation of DNA loops. One main aim of this paper is to clearly show how the interaction between DNA loops results in the differences between free energies at different states. For this, we transform this issue into that of the free energy dissipation defined as the entropy production rate. For clarity, we first consider a simple sub-block of two states OFF1 and OFF2 in Fig. 1. Denote by F_1 and F_2 the free energies that the gene is at OFF1 and OFF2 states, respectively. Then, according to Ref. [31], we know that the ratio between two transition rates, $\lambda_{12}/\lambda_{21}$, is proportional to $e^{-\beta\Delta F_1}$, i.e., $\lambda_{12}/\lambda_{21} \propto e^{-\beta\Delta F_1}$, where $\Delta F_1 = F_2 - F_1$ represents the difference between free energies, F_1 and F_2 (i.e., the change in free energy), and $\beta = 1/(k_B T)$ is a composite parameter of the Boltzmann constant and temperature (without loss of generality, we set $\beta = 1$ in our analysis). We can show that the free energy consumption for the OFF1–OFF2 block is given by $\dot{\omega}_1 = \frac{d}{dt} \omega_1 = J \ln \frac{\lambda_{12}}{\lambda_{21}}$, where J is a constant, and will be specified later. Thus, $\dot{\omega}_1 = h_1 - J\Delta F_1$, where h_1 is also a constant. In other words, the relation between $\dot{\omega}_1$ and ΔF_1 is linear, and the difference between the free energy consumption rate and the free energy difference is determined by a constant factor h_1 and a constant multiplier J , where the former constant is interest of this paper whereas the latter constant depends usually on the hydrolyzes of ATPs (energetic molecules)^{45–48}. Similarly, if we denote by F_3 and F_4 the free energies that the gene is at ON1 and ON2 states respectively, and by $\Delta F_2 = F_3 - F_2$, $\Delta F_3 = F_4 - F_3$ and $\Delta F_4 = F_1 - F_4$ the differences between free energies of the system's states, then we have $\frac{\lambda_{23}}{\lambda_{32}} \propto e^{-\beta\Delta F_2}$, $\frac{\lambda_{34}}{\lambda_{43}} \propto e^{-\beta\Delta F_3}$ and $\frac{\lambda_{41}}{\lambda_{14}} \propto e^{-\beta\Delta F_4}$. To help the reader's understanding, we take the nested loops structure as an example to show the changes of free energy in each state, referring to Fig. 2B. For this figure, we give interpretations below. Since free energy is a measuring index in statistical physics, and its size is unknown in most cases. Without loss of generality, we may assume that the free energy

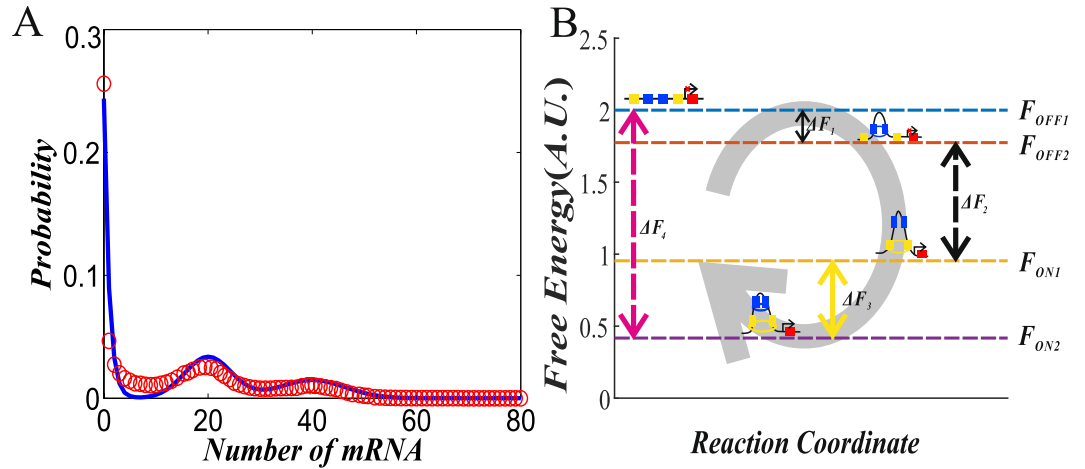


Figure 2. (A) Comparison probability distributions between analytical (solid line) and numerical (empty circles) results, where parameter values are set as $\lambda_{12} = 0.4$, $\lambda_{21} = 0.2$, $\lambda_{23} = 0.2$, $\lambda_{32} = 0.3$, $\lambda_{34} = 0.1$, $\lambda_{43} = 0.1$, $\lambda_{41} = 0.5$, $\lambda_{14} = 0.5$, $\mu_1 = 20$, $\mu_2 = 40$ and $\delta = 1$. (B) Free energies at 4 states of nested loops structure, where the free energy of the OFF1 state is set as 2 and energy differences between the states (denoted by ΔF_i , $i = 1, 2, 3, 4$) are indicated. Parameter values are set as $d_1 = 500$, $d_2 = 300$ and $\lambda_{12} = \lambda_{32} = \lambda_{34} = \lambda_{41} = 0.3$.

of the OFF1 state is 2 (due to the setting of $\beta = 1$). By calculating the free energy difference between different states, one can judge the energy cost of switching between these states. From Fig. 2B, we observe that increasing the number of promoter states can reduce free energy, implying that the formation of DNA loop needs free energy.

Furthermore, the energy dissipation rates, $\dot{\omega}_i$, can also be expressed using the differences between free energies, that is, $\dot{\omega}_i = h_i - J\Delta F_i$, where $i = 1, 2, 3, 4$.

Finally, we set $\Delta F = \Delta F_1 + \Delta F_2 + \Delta F_3 + \Delta F_4$, which represents the change in free energy for the cyclic promoter of the gene shown in Fig. 1C or D. Note that $\frac{\lambda_{12}\lambda_{23}\lambda_{34}\lambda_{41}}{\lambda_{43}\lambda_{32}\lambda_{21}\lambda_{14}} \propto e^{-\beta\Delta F}$ or $\ln\left(\frac{\lambda_{12}\lambda_{23}\lambda_{34}\lambda_{41}}{\lambda_{43}\lambda_{32}\lambda_{21}\lambda_{14}}\right) = h - \beta\Delta F$, where h is a constant depending on both the Boltzmann constant and temperature. Thus, we obtain the following expression for the relationship between the dissipation rate of free energy for the gene promoter and the difference between the corresponding free energies

$$\dot{W} = h - J\Delta F \tag{6}$$

This establishes the linear relationship between the free energy difference ΔF and the free energy dissipation rate \dot{W} . Recall that the size of the difference between free energies relies on hydrolyzes of energetic molecules such as ATPs^{45–48}. Therefore, studying the free energy consumption rate in a system of stochastic gene expression can help us understand the roles of regulation factors or processes such as the interaction between DNA loops in controlling the expression level. In this paper, we are more interested in mean free energy consumption rate (or simply “mean energy”), which is defined as the ratio of the energy dissipation rate over the mean mRNA. According to the formula (6), we know that the mean free energy consumption rate is proportional to the free energy required per mRNA. We will apply the minimal energy principle to speculate the mechanism of the interaction between DNA loops.

An effective method for calculating the free energy cost of the whole system. Based on the above results, here we provide an effective method for calculating the free energy cost of the entire gene expression system. First, we introduce 4 logic variables, $\tilde{x}_1, \tilde{x}_2, \tilde{x}_3$ and \tilde{x}_4 , to represent 4 promoter states, where every \tilde{x}_i takes only 0 or 1, i.e., $\tilde{x}_i \in \{0, 1\}$. Since the gene must be and is only at one of 4 states, we have the conservative condition: $\tilde{x}_1 + \tilde{x}_2 + \tilde{x}_3 + \tilde{x}_4 = 1$. For a fixed set of $\tilde{x}_1, \tilde{x}_2, \tilde{x}_3$ and \tilde{x}_4 , the Fokker-Plank equation for y (where $dy/dt = \mu_1x_4 + \mu_2x_3 - \delta y \equiv F$) can be approximated as

$$\frac{\partial P(y, t)}{\partial t} = -\frac{\partial}{\partial y} \left(FP - \frac{1}{2} \frac{\partial}{\partial y} (\Phi P) \right) \tag{7}$$

where $\Phi = \mu_1x_4^s + \mu_2x_3^s + \delta y^s = \frac{\mu_2C + \mu_1D}{E} + \delta y^s = 2 \frac{\mu_2C + \mu_1D}{E}$

According to refs^{37,49–51}, we know that the dissipation rate of free energy can in general be expressed as

$$\dot{W} \equiv \frac{dW}{dt} = \sum_{A,B} (J_{A \rightarrow B} - J_{B \rightarrow A}) \ln \frac{J_{A \rightarrow B}}{J_{B \rightarrow A}} \tag{8}$$

where A and B represent the microscopic states of the underlying system, and $J_{\sigma \rightarrow \sigma'}$ represents the transition probability from state σ to state σ' . In our case, A and B represent the states specified by $A = (\tilde{x}_1, \tilde{x}_2, \tilde{x}_3, \tilde{x}_4, y)$ and

$B = (\tilde{x}_1, \tilde{x}_2, \tilde{x}_3, \tilde{x}_4, y + \Delta y)$, where the absolute $|\Delta y|$ is infinitesimal, and y is a continuous variable. Moreover, the following decomposition holds

$$\sum_{A,B} (J_{A \rightarrow B} - J_{B \rightarrow A}) \ln \frac{J_{A \rightarrow B}}{J_{B \rightarrow A}} = \sum_{\tilde{x}_1, \tilde{x}_2, \tilde{x}_3, \tilde{x}_4} \int (J_{A \rightarrow B} - J_{B \rightarrow A}) \ln \frac{J_{A \rightarrow B}}{J_{B \rightarrow A}} dy + \sum_{\tilde{x}_1, \tilde{x}_2, \tilde{x}_3, \tilde{x}_4} \dot{\omega}_y \quad (9)$$

where the first term on the right-hand side of Eq. (9) represents the free energy dissipation along the hyperplane $\tilde{x}_1 + \tilde{x}_2 + \tilde{x}_3 + \tilde{x}_4 = 1$ in the state space whereas the second term represents the free energy dissipation along the y -direction. Thus,

$$\dot{W} = \dot{W}_p + \dot{W}_y \quad (10)$$

where

$$\dot{W}_p = \sum_{\tilde{x}_1, \tilde{x}_2, \tilde{x}_3, \tilde{x}_4} \int (J_{A \rightarrow B} - J_{B \rightarrow A}) \ln \frac{J_{A \rightarrow B}}{J_{B \rightarrow A}} dy \quad (10a)$$

$$\dot{W}_y = \sum_{\tilde{x}_1, \tilde{x}_2, \tilde{x}_3, \tilde{x}_4} \dot{\omega}_y \quad (10b)$$

The left task is to specify the expressions of \dot{W}_p and \dot{W}_y . In the Supporting Online Material of this paper, we have showed

$$\dot{W}_p = \frac{\lambda_{12}\lambda_{23}\lambda_{34}\lambda_{41} - \lambda_{43}\lambda_{32}\lambda_{21}\lambda_{14}}{E} \ln \frac{\lambda_{12}\lambda_{23}\lambda_{34}\lambda_{41}}{\lambda_{43}\lambda_{32}\lambda_{21}\lambda_{14}} \equiv J \ln q \quad (11)$$

which is just the energy dissipation of a circulation per unit time (see refs^{38,52,53}), where $J = \frac{\lambda_{12}\lambda_{23}\lambda_{34}\lambda_{41} - \lambda_{43}\lambda_{32}\lambda_{21}\lambda_{14}}{E}$ and $q = \frac{\lambda_{12}\lambda_{23}\lambda_{34}\lambda_{41}}{\lambda_{43}\lambda_{32}\lambda_{21}\lambda_{14}}$. Regarding to the calculation formula for \dot{W}_y , we can show (seeing the Supporting Online Material for details)

$$\begin{aligned} \dot{W} &= \left(\frac{\lambda_{12}\lambda_{23}\lambda_{34}\lambda_{41} - \lambda_{43}\lambda_{32}\lambda_{21}\lambda_{14}}{E} \right) \ln \frac{\lambda_{12}\lambda_{23}\lambda_{34}\lambda_{41}}{\lambda_{43}\lambda_{32}\lambda_{21}\lambda_{14}} \\ &+ \frac{2}{\Phi} \left[\frac{2A + 2B}{E} + \frac{\delta^2(C\sigma_2^2 + D\sigma_1^2)}{E} \right] + \frac{2\delta(A + B - C - D)}{E} \\ &+ \frac{\Phi}{2E} \left[(A + B)\delta^2 + \frac{C}{\sigma_2^2} + \frac{D}{\sigma_1^2} \right] \end{aligned} \quad (12)$$

Results

Influence of the DNA loop length on free energy consumption. As is well known, whether two DNA regulatory elements form a loop depends on the distance between them along the DNA line, and that this distance in turn can affect gene expression and further cell-to-cell variability^{18,19}. Here, we investigate how DNA loop lengths in the possible structures of two interacting DNA loops impact the free energy dissipation rate (or simply “energy”, which is equal to the entropy production rate. See Methods for details), the mean mRNA, and the mean free energy consumption rate (i.e., “mean energy”). For clarity, we change a DNA loop length (d_1) while keeping another DNA loop length (d_2) fixed. The numerical results are shown in Fig. 3, where the value range of d_1 is from the experimental datas (12).

From Fig. 3, we observe that for each of three fundamental structures (nested, side-by-side and alternating), there is an optimal loop length (d_1) such that the energy dissipation rate reaches a maximum. Moreover, in contrast to the side-by-side structure, the nested structure has a larger energy dissipation rate (i.e., consuming more free energy) while the alternating structure has a smaller energy dissipation rate, referring to Fig. 3A. On the other hand, for each structure, the mean mRNA expression level is all a monotonically decreasing function of d_1 , but is higher for the side-by-side structure than for the alternating structure but lower than for the nested structure, referring to Fig. 3B. Although the energy dissipation rate is not a monotonic function of d_1 (Fig. 1A), the mean energy dissipation rate is, referring to Fig. 3C. Interestingly, the order for three curves shown in Fig. 3A,B currently becomes opposite in the case of Fig. 3C (i.e., the case of mean energy consumption rate). Figure 3 indicates that in all the three structures, the nested structure produces most mRNAs and consumes least mean free energy (meaning that generating one mRNA dissipates least free energy). By contrast, the alternating structure produces fewest mRNAs and consumes most mean free energy (meaning that generating one mRNA dissipates most energy). Thus, we conclude that the nested structure performs best in all the three structures from the viewpoint of energy consumption.

Note that the further the distance between regulatory elements is, the more difficult they form a DNA loop or the smaller the DNA looping rate becomes. Thus, we also conclude that the faster the DNA looping becomes, the less mean free energy is dissipated and the more mean mRNAs is produced, whereas the slower the DNA looping becomes, the more mean free energy is consumed and the fewer mean mRNAs is produced.

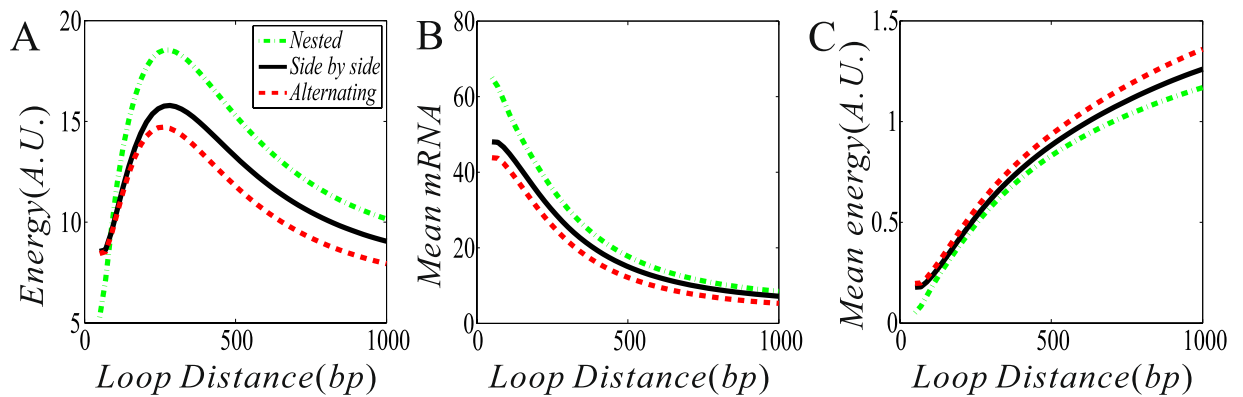


Figure 3. The effect of a DNA loop length on the free energy dissipation rate (energy) (no unit), the mRNA expression level, and the free energy dissipation rate (mean energy) (no unit). Where parameter values in all the cases are set $\mu_1 = 40$, $\mu_2 = 80$, $\delta = 1$, $\gamma = 0$, $d_2 = 300$, $\lambda_{12} = \lambda_{32} = \lambda_{34} = \lambda_{41} = 0.3$, and $d_1 \in (50, 1000)$.

Influence of the interaction between DNA loops on free energy consumption. As is well known, the interaction between DNA loops would be complex^{2,4,5,12,13}. Many questions, e.g., how the loop-loop interactions including communication forms affect gene expression and how DNA looping consumes free energy, remain elusive. In the last subsection, we have shown that if one DNA loop formation promotes another DNA loop formation (e.g., in the nested structure), then free energy can be saved (i.e., the promotion reduces free energy consumption). By contrast, if one DNA loop formation suppresses another DNA loop formation (e.g. in the alternating structure), then more free energy is consumed (i.e., the suppression increases free energy consumption). In this subsection, we consider another mode of the interaction between two DNA loops, i.e., one DNA loop is assumed to influence the transcriptional rate of another DNA loop.

To help the reader understand the results obtained in this subsection, let us recall our assumptions (see details in Material and Methods). The blue DNA loop formed by a pair of regulatory elements interacts with the yellow DNA loop formed by a pair of enhancer and promoter, and that the gene is not expressed in the former case but is expressed in the latter case. The former loop may affect the latter loop in two ways: promotion and suppression. Specifically, if the blue loop facilitates the formation of the yellow loop, i.e., if the former enlarges the looping rate of the latter or the transcription rate, then we call the corresponding case as promotion. If the blue loop reduces the looping rate of the yellow loop or the transcription rate, then we call the corresponding case as suppression. In the above subsection, we have analyzed the effect of looping rate on free energy consumption. In this subsection, we try to answer the question of how the promotion or the suppression impacts free energy consumption, mean mRNA expression and mean free energy dissipation. By numerical analysis, we find some universal phenomena (see the following contents). For clarity, we distinguish the following into two cases: one DNA loop promotes the other DNA loop (more precisely, the former enhances transcription of the latter); one DNA loop suppresses the other DNA loop. In the following, μ_1 represents the transcription rate at ON1 (in this case, the yellow loop is formed but the blue loop is not formed, so μ_1 may be understood as a fundamental transcription rate); μ_2 represents the transcription rate at ON2 (in this case, both the yellow loop and the blue loop are formed, so μ_2 may be understood as a regulated transcription rate). Note that $\mu_1 < \mu_2$ means that the blue loop formation promotes the yellow loop formation, whereas $\mu_1 > \mu_2$ means that the former prohibits the latter. Note that the size of μ_2 can represent the interaction degree.

First, consider the case that one DNA loop enhances the expression of the other DNA loop, i.e., consider the case of $\mu_1 \mu_2$. From Fig. 4A, we observe that if the transcription rate μ_1 is fixed, then the dissipation rate of free energy is a monotonically increasing function of the transcription rate μ_2 , whichever the structure of two DNA loops. However, there are differences in the amount of energy consumption among three fundamental structures. Specifically, in contrast to the side-by-side structure, the nested structure consumes more free energy due to the increase of the mRNA mean whereas the alternating structure consumes less free energy due to the decrease of the mRNA mean, independent of μ_2 that must be larger than μ_1 due to promotion. Results in the case of mean mRNA expression analysis is fundamentally the same as those in the case of free energy consumption rate analysis, referring to Fig. 4B. However, the change tendency and the order of three curves in the case of mean free energy dissipation are completely different from the case without making average, referring to Fig. 4C. Precisely, the mean free energy dissipation rate is a monotonically decreasing function of the transcription rate μ_2 , and the order of three curves is that the curve for the nested structure is below the curve for the side-by-side structure, which is below the curve for the alternating structure.

From Fig. 4, we can conclude some results as following:

- (1) Regardless of the structure of the interaction between two DNA loops, if μ_2 increases (i.e., the promotion interaction is enhanced), then the mean RNA level will increase (following the increase in free energy consumption) and the mean energy consumption (i.e., the free energy consumed by the production of one mRNA) decreases. As such, we speculate that the promotion-type interaction may save free energy.
- (2) In contrast to the side-by-side structure, the nested structure consumes less mean free energy while the

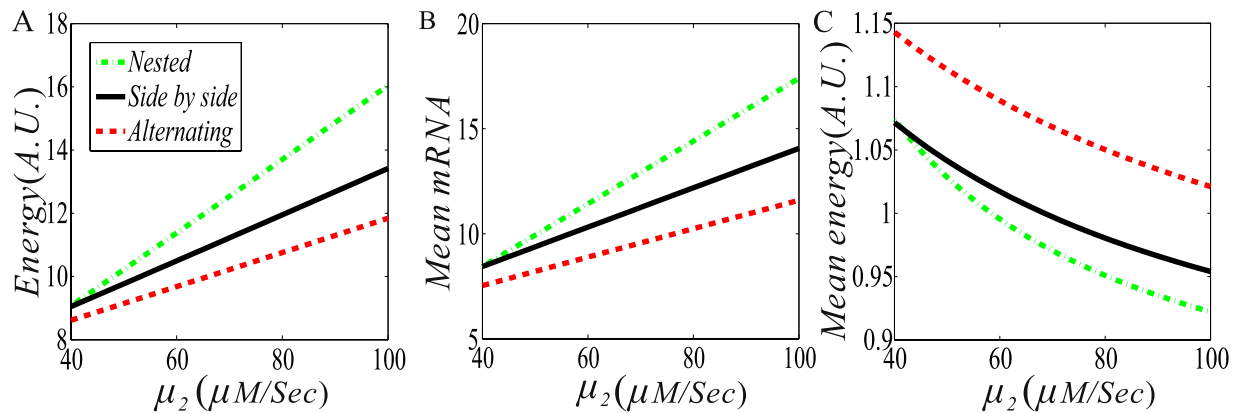


Figure 4. The effect of the promotion (i.e., $\mu_1 < \mu_2$) of two DNA loops on free energy dissipation and mRNA expression level. **(A)** The dissipation rate of free energy vs the transcription rate (μ_2); **(B)** The mean mRNA expression level vs μ_2 ; and **(C)** The mean free energy consumption vs μ_2 . In all the cases, parameter values are set as $\mu_1 = 40$, $\mu_2 \in (40, 100)$, $\delta = 1$, $\gamma = 0$, $d_2 = 300$, $d_1 = 500$, $\lambda_{12} = \lambda_{32} = \lambda_{34} = \lambda_{41} = 0.3$.

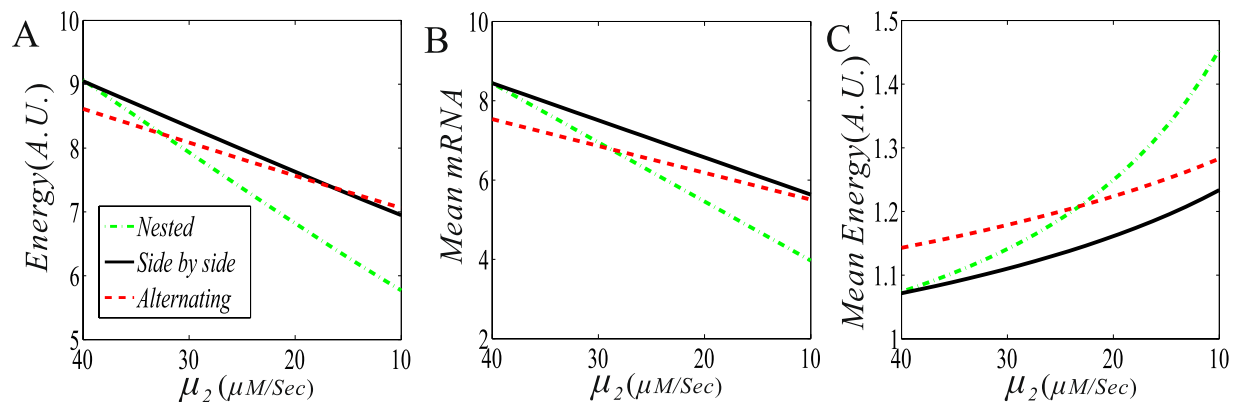


Figure 5. The effect of the suppression (i.e., $\mu_1 > \mu_2$) of two DNA loops on free energy dissipation and mRNA expression level. **(A)** The dissipation rate of free energy vs the transcription rate (μ_2); **(B)** The mean mRNA expression level vs μ_2 ; and **(C)** The mean free energy consumption vs μ_2 . In all the cases, except for $\mu_2 \in (10, 40)$, other parameter values are set as the same as in Fig. 4.

alternating structure consumes more free mean energy if one DNA loop facilitates the other DNA loop. The result is opposite otherwise.

Then, consider the case that one DNA loop represses the expression of the other DNA loop, i.e., consider the case of $\mu_2 < \mu_1$. The numerical results are shown in Fig. 5. From Fig. 5A, we observe that in contrast to the side-by-side structure, both the nested and the alternating structures always consume less free energy since the corresponding curves are below the curve for the side-by-side loops. For a smaller μ_2 of $\mu_2 < \mu_1$ (e.g., $\mu_2 < 33$ if $\mu_1 = 40$ is set), the nested structure consumes less free energy than the alternating structure, whereas for a larger μ_2 of $\mu_2 < \mu_1$, the former consumes more free energy. The dependence of the mean mRNA level on μ_2 has the similar change tendency to that of the free energy dissipation rate, but the critical value of μ_2 for the cross point of two curves currently become smaller, referring to Fig. 5B. Then, we analyze Fig. 5C. We observe that the mean free energy consumption is a monotonically decreasing function of μ_2 , independent of the structure of two DNA loops. In contrast to the side-by-side structure, both the nested and the alternating structures consume more mean free energy since the corresponding curves are beyond the curve for the side-by-side loops. For a smaller μ_2 of $\mu_2 < \mu_1$ (e.g., $\mu_2 < 23$ if $\mu_1 = 40$ is set), the nested structure consumes more free energy than the alternating structure, whereas for a larger μ_2 of $\mu_2 < \mu_1$, the former consumes less mean free energy than the latter.

From Fig. 5, we can conclude that if the blue DNA loop suppresses the yellow DNA loop (meaning that the former reduces the transcription rate of the latter), then both the free energy dissipation and the mean mRNA level are reduced but the mean free energy consumption is increased, regardless of the structure between two interacting DNA loops. Moreover, the smaller the μ_2 is (or the stronger the suppression is), such a reduction or increase becomes more apparently. Simply speaking, the suppression leads to the increasing of the mean free energy consumption (meaning that the production of one mRNA needs to consume more free energy). Therefore, we speculate that the suppression-type interaction needs to dissipate more free energy.

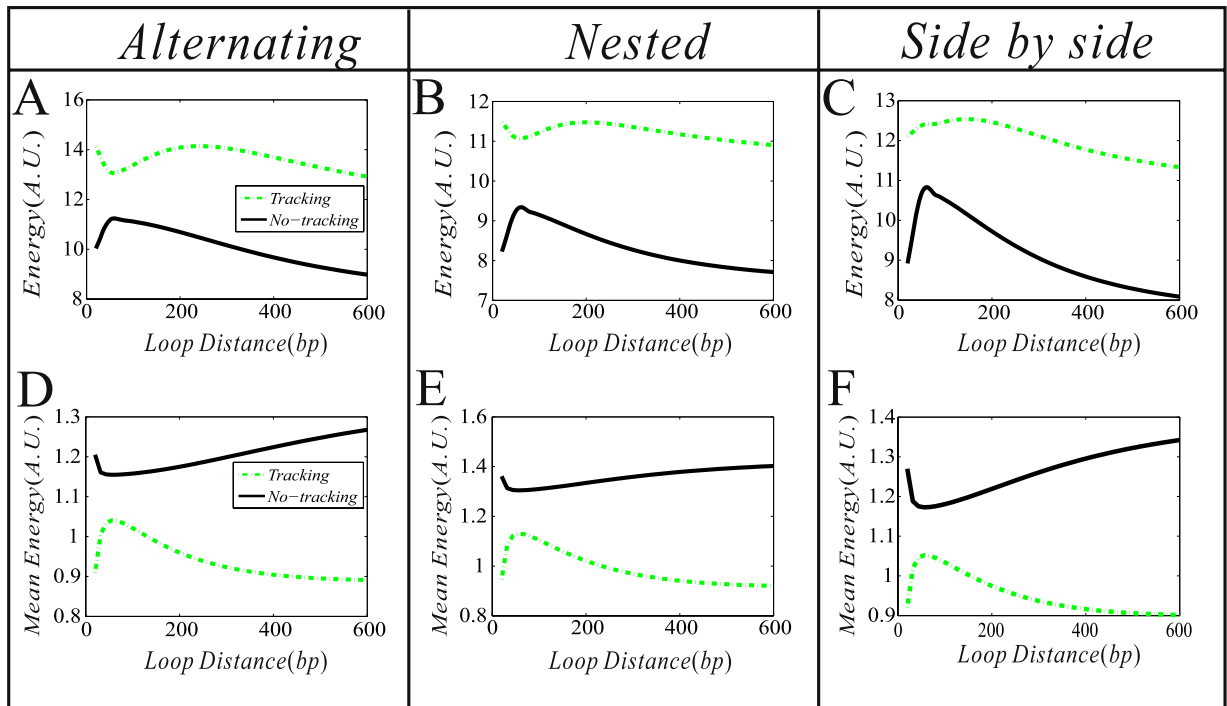


Figure 6. The dependences of energy dissipation rate and the mean energy dissipation rate on the length of the blue loop. Dashed lines corresponding to the facilitated-tracking model ($\gamma \neq 0$) whereas solid lines to the directing looping model ($\gamma = 0$). In all the cases, parameter values are set as $\mu_1 = 40$, $\mu_2 = 80$, $\delta = 1$, $\gamma = 0.1$, $d_2 \in (10, 600)$, $d_1 = 1000$, $\lambda_{12} = \lambda_{32} = \lambda_{34} = \lambda_{41} = 0.3$.

By comparing Figs 4 and 5, we obtain a universal conclusion, that is, the promotion of two loops dissipates more free energy but less mean free energy whereas the suppression consumes less free energy but more mean free energy, whichever the structure of two DNA loops.

Influence of the communication form between regulatory elements on free energy consumption. In biology, a long-term debate is which of direct looping model and facilitated-tracking model is more reasonable. Here, we try to give an answer to this question from the viewpoint of free energy dissipation.

Before that, we first introduce a parameter to quantify the effect of the communication way between DNA regulatory elements on gene expression. Imagine a DNA loop as a string with a fixed length, two ends of which represent looping elements (e.g., a pair of Su and Hw, a pair of Anchor and CTCF). If one element slides along this string (here we only consider the sliding of one element in the blue loop), then this will affect the range that the enhancer and the promoter form the yellow loop. Thus, the tracking mechanism leads to the increase in looping rates. Specifically, if the looping rates of the yellow and blue loops are denoted by a pair of $\tilde{\lambda}_{14}$ and $\tilde{\lambda}_{12}$ in the case that the facilitated-tracking mechanism exists, and their natural looping rates by another pair of λ_{14} and λ_{12} , then the relationships between these pairs can be expressed as

$$\tilde{\lambda}_{14} = \lambda_{14} + \Delta_1, \quad \tilde{\lambda}_{12} = \lambda_{12} + \Delta_2 \quad (13)$$

where Δ_i represents the differences between the two cases. Similarly, a pair of λ_{23} and λ_{34} need to be modified. Note that for the facilitated-tracking mechanism, a longer DNA loop leads to a wider range for one regulatory element to track another regulatory element, implying $\Delta \sim d$ or $\Delta = rd$, where r is a nonnegative parameter. Thus, no tracking or direct looping corresponds to $r = 0$, whereas the facilitated-tracking model corresponds to $r \neq 0$. In ref.¹⁸, the parameter r is called as the tracking ratio, which can be understood as the probability that the enhancer and the promoter track to each other along the DNA line.

Now, we examine the effect of the communication way between loop elements on free energy dissipation and on the mean free energy dissipation. First, investigate the dependences of free energy dissipation and the mean free energy dissipation on the length of the blue loop (d_2) in both cases of tracking (corresponding to $r \neq 0$) and of no tracking (corresponding to $r = 0$). The numerical results are shown in Fig. 6.

From Fig. 6, we first observe that more free energy is consumed in the case of tracking (e.g., $r = 0.1$) than in the case of no tracking (meaning $r = 0$) in all the three structures (comparing dash lines with solid lines in Fig. 6A–C). However, less mean free energy is dissipated in the case of tracking than in the case of no tracking (comparing dash lines with solid lines in Fig. 6D–F). This indicates that from the viewpoint of average, the facilitated-tracking mechanism is better in free energy dissipation than the direct looping mechanism. In addition, it implies that the former communication mechanism facilitates the mRNA expression than the latter communication mechanism since the mean free energy consumption is reduced. The above observation is the main result of this subsection.

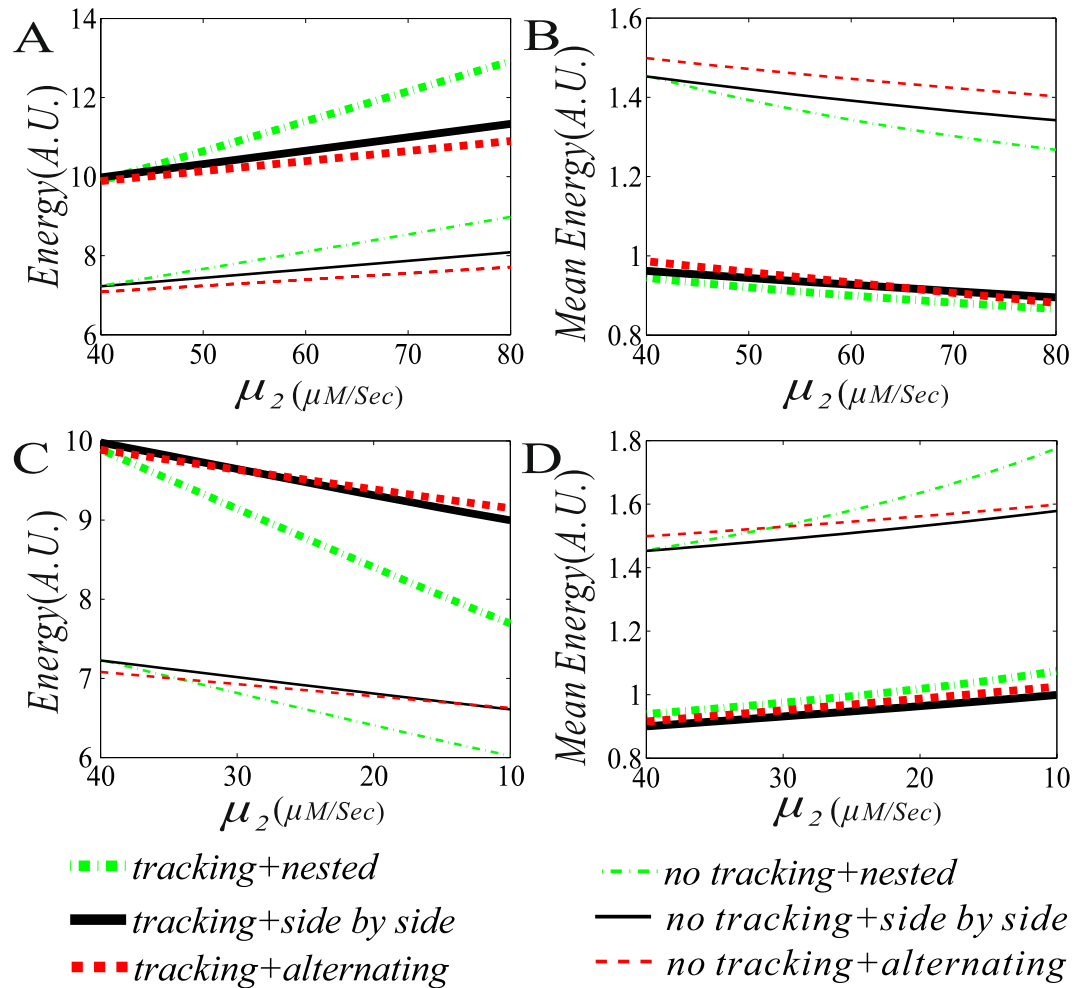


Figure 7. Comparison between the results for tracking and those for no tracking. (A) and (B), the free energy consumption rate/mean free energy consumption rate vs the transcription rate μ_2 , where the blue DNA loop is assumed to enhance the expression of the yellow DNA loop, $\mu_1 = 40$, and $\mu_2 \in (40, 80)$; (C) and (D), the free energy consumption rate/mean free energy consumption rate vs the transcription rate μ_2 , where the blue DNA loop is assumed to repress the expression of the yellow DNA loop, $\mu_1 = 40$, and $\mu_2 \in (10, 40)$. In all the cases, the other parameter values are set as $\delta = 1$, $\gamma = 0.1$, $d_2 = 600$, $d_1 = 1000$, and $\lambda_{12} = \lambda_{32} = \lambda_{34} = \lambda_{41} = 0.3$.

Next, we show the effects of promotion (i.e., $\mu_2 > \mu_1$) and suppression (i.e., $\mu_2 < \mu_1$) on free energy consumption and on its mean in the case that the facilitated-tracking mechanism is considered. For this, we compare the results in the case of tracking with those in the case of no tracking (i.e., direct looping), referring to Fig. 7. We observe from this figure that the more free energy is dissipated in the case of tracking (corresponding to thick curves in Fig. 7A,C) than in the case of no tracking (corresponding to thin curves in Fig. 7A,C). In contrast, the less mean free energy is dissipated in the case of tracking (corresponding to thick curves in Fig. 7B,D) than in the case of no tracking (corresponding to thin curves in Fig. 7B,D). These results are independent of the way of the interaction between two DNA loops.

In order to show the global effect of promotion (i.e., $\mu_1 < \mu_2$)/suppression (i.e., $\mu_1 > \mu_2$) and the tracking ratio on both the free energy dissipation and the mean free energy dissipation, we further plot Fig. 8, a three-dimensional pseudo diagram. From the Figure, we clearly observe that the larger both the μ_2 of $\mu_1 < \mu_2$ and the tracking ratio (r) are, the more is the total free energy is consumed (see Fig. 8A), but the less is the mean total free energy consumed (see Fig. 8B) since the mean mRNA level is increased. In contrast, the smaller the μ_2 of $\mu_2 < \mu_1$ is and the tracking ratio (r) are, the less is the total free energy is consumed (see Fig. 8C), but the more is the mean total free energy consumed (see Fig. 8D) since the mean mRNA level is decreased. Thus, the results shown in Fig. 8 are in agreement with those shown in Fig. 7.

The above analysis indicates that the facilitated-tracking mechanism always reduces the mean free energy dissipation in contrast to the direct-looping mechanism. In addition, we have shown that the promotion-type interaction between two DNA loops may save free energy in contrast to the suppression-type interaction. Thus, according to the minimal energy principle, we speculate that the promotion-type interaction between two DNA loops plus the facilitated-tracking looping mechanism is the most possible way utilized by live organisms. Related biological reasons for this speculation are stated as follow.

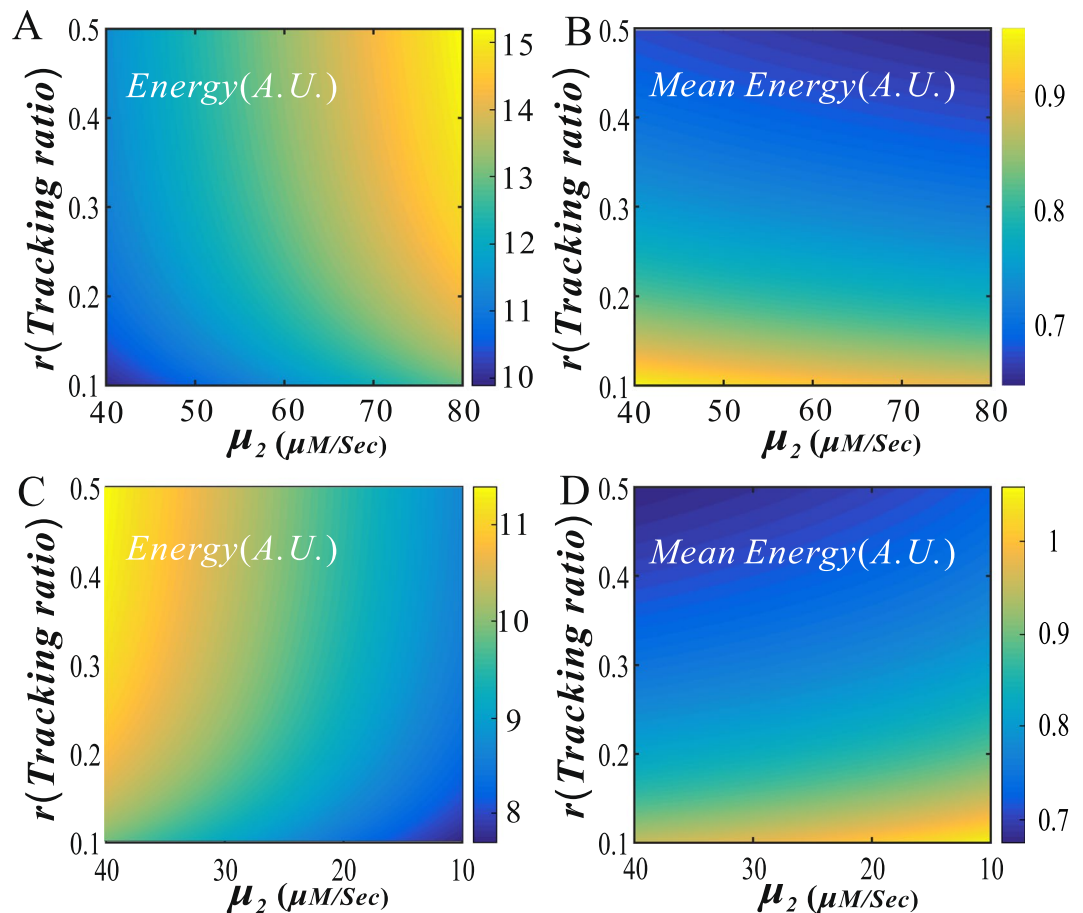


Figure 8. Three-dimensional pseudo diagrams for dependences of free energy dissipation rate/the mean free energy dissipation rate on both the transcriptional rate and the tracking ratio. (A) and (B) the blue loop promotes the transcription rate of the yellow loop, where parameters are set as $\mu_1 = 40$, $\mu_2 \in (40, 80)$, $\delta = 1$, $\gamma \in (0.1, 0.5)$, $d_2 = 600$, $d_1 = 1000$, $\lambda_{12} = \lambda_{32} = \lambda_{34} = \lambda_{41} = 0.3$, $k_1 = 40 e^{-0.05 d_1} + 1$, $k_2 = 4$, (C) and (D) the blue loop suppresses the transcription rate of the yellow loop, except for $\mu_2 \in (10, 40)$, other parameter values are set as the same as in (A).

First, two mainstream communication forms between DNA regulatory elements: direct looping and facilitated-tracking looping, exist extensively in real biological systems. For example, experimental data or evidence support the mechanism of facilitated-tracking looping between enhancer and promoter⁵⁴, whereas other experimental data or evidence support the mechanism of direct looping between enhancer and promoter^{55–57}. Second, that live organisms adopt which of the two mechanisms is a long-term debate in biology¹⁵. Third, experiments found that one DNA loop may influence the expression of another DNA loops^{22,58}. However, how they influence each other is not only unclear but also difficult to measure by an experimental method. Here we apply the minimal energy principle to give a positive answer to this issue, as stated above.

Discussion

Non-equilibrium mechanisms play important roles in many biological processes ranging from the concentration gradients that cells establish both with their environments and within themselves to chaperone-assisted protein folding and to gene expression. These non-equilibrium processes are essential for life as Ahsendorf, *et al.*, ever pointed out that “we are only at equilibrium when we are dead”⁵⁹. From the viewpoint of thermodynamics, non-equilibrium processes necessarily consume energy^{38,39}. For example, the formation of a single DNA loop consumes about 9 kcal/mol but the corresponding energetic cost can be overcompensated by the interaction energy with transcription factors of some type that maintain the loop⁶⁰. From the perspective of information theory, the entropy production rate is precisely the amount of energy consumption⁵³. There is no energy consumption for detailed-balance systems but there is energy consumption for non-equilibrium steady-state systems⁵². However, quantitative analysis of free energy dissipation in biological systems, in particular in those of gene expression regulation, is nontrivial due to complexity of the involved biochemical processes. In this paper, by analyzing the free-energy costs of DNA looping and the interaction between interacting DNA loops, we have found universal results, e.g., whichever the structure of two loops (nested, side-by-side or alternating), the promotion of one DNA loop to another DNA loop (including increasing the looping rate and the transcription rate) always consumes less mean free energy whereas the suppression has the just opposite effect. More interestingly, we have shown that in contrast to the mechanism of direct looping between regulatory elements, the facilitated-tracking

mechanism consumes less mean free energy but can enhance the mean mRNA expression. This result justifies the facilitated-tracking hypothesis, a long-standing debate in biology.

We have analyzed the free-energy costs in three fundamental structures of DNA-looping interactions (alternating loops, nested loops, and side-by-side loops), but enhancers and promoters may be connected in a highly complex network of DNA-looping interactions^{61–63}, remarkably in eukaryotic cells. Since, many questions, e.g., at which step during gene activation, various nucleoprotein complexes assemble at distant enhancers, and how these complexes then contribute to promoter accessibility, the preinitiation complex recruitment and/or assembly, and transcription initiation and elongation, have been unsolved, the mechanisms for the energetic cost of gene expression have not been completely elucidated. In addition, enhancers have been shown to have a role in the preinitiation complex recruitment at target promoters^{64–66}, the removal of proteasome complexes at promoters⁶⁷, the generation of intra-chromosomal loops between regulatory regions⁶⁸, and the regulation of elongation^{69,70}; Enhancers are also involved in the removal of repressive histone modifications^{71–74}, suggesting that they also contribute to the delivery of enzymes that regulate histone modifications^{75,76}. In a word, enhancers in eukaryotic genomes can be many hundreds of kilobases away from the promoter they regulate^{76, 77,78}, and the intervening DNA can contain other promoters and other enhancers^{61,79,80}. All these complex cases would greatly complicate the investigation of the energetic cost in gene expression, and it is needed to develop new models and computational methods. However, our model has plasticity in many aspects, e.g., it can easily incorporate three main factors: connection pattern, distance between regulatory elements and communication form, which altogether can characterize interactions between chromatin loops.

From the perspective of applications, our method would provide a paradigm for analyzing the free-energy cost in gene expression involving complex regulatory processes. First, according to our proposed map method, we can map the topologies for the interactions among arbitrary DNA loops into a multistep model of gene expression, where DNA loop lengths (along the DNA lines) and other rates quantifying elaborate processes such as tracking between regulatory elements, energy-dependent chromatin remodeling, are easily incorporated into transition rates between promoter states, as done in this paper. This mapping is a key for one to investigate the corresponding energetic cost. Then, recall that the Gibbs energy is defined as

$$W = -k \int P(\mathbf{x}; t) \ln P(\mathbf{x}; t) d\mathbf{x} \quad (14)$$

where $k = k_B T$ with k_B being a Boltzmann constant and T being temperature, and the entropy production rate that quantifies energy dissipation is defined as $\dot{W} = dW/dt$. To calculate \dot{W} , it is required to know the joint probability distribution $P(\mathbf{x}; t)$. However, even if we know that the Fokker-Planck equation for the underlying biochemical system is given by

$$\frac{\partial P}{\partial t} = - \sum_{i=1}^{n+1} \frac{\partial}{\partial x_i} \left(F_i P - \frac{1}{2} \sum_{j=1}^{n+1} \frac{\partial}{\partial x_j} (\Phi_{ij} P) \right) \equiv - \sum_{i=1}^{n+1} \frac{\partial J_i}{\partial x_i} \quad (15)$$

where x_i quantifies promoter state i (e.g., representing the proportion of the DNA number at this state divided by the total DNA number), F_i represents dynamics of state i subjected to noise with the intensity Φ_{ij} ($1 \leq i \leq n$), and $y \equiv x_{n+1}$ represents gene production (mRNA or protein), it is very difficult to derive the expression of \dot{W} . In fact, we can only perform the formal calculation in this case (as done in most of the existing references):

$$\begin{aligned} \dot{W} &= -k \int [\ln P(\mathbf{x}; t) + 1] \frac{\partial P(\mathbf{x}; t)}{\partial t} d\mathbf{x} \\ &= k \sum_{i=1}^{n+1} \int [\ln P(\mathbf{x}; t) + 1] \frac{\partial J_i}{\partial x_i} d\mathbf{x} \\ &= -k \sum_{i=1}^{n+1} \int \frac{J_i}{P} \frac{\partial P}{\partial x_i} d\mathbf{x} \\ &= \frac{k}{2} \sum_{i=1}^{n+1} \sum_{j=1}^{n+1} \int \frac{1}{P} \frac{\partial P}{\partial x_i} \frac{\partial (\Phi_{ij} P)}{\partial x_j} d\mathbf{x} - k \sum_{i=1}^{n+1} \int F_i \frac{\partial P}{\partial x_i} d\mathbf{x} \end{aligned}$$

In spite of this, finding the distribution $P(\mathbf{x}; t)$ is another key for one to investigate the energetic cost and in general difficult, in particular in the cases that many complex processes associated with gene expression are considered. In this paper, we have proposed a simple approach to find $P(\mathbf{x}; t)$, which is based on the particular structure mapped from a complex network for the interactions among chromatin loops as well as the probability's sum rule for independent events.

Finally, it should be pointed out that regulation (including the formation of DNA loops and the interaction between loops) is classically approached with thermodynamic methods^{36–39}. We have shown that our model can be expressed in energetic terms and constitute a generalization of these approaches by extending the promoter structure, the range of systems that can be represented (i.e., including energy consuming systems such as eukaryotic promoters), and the type of metrics that can be predicted (i.e., including measures of dynamic and stochastic properties). The usual thermodynamic formulation of cooperative and competitive association/dissociation of transcription factors (TFs)³¹ is equivalent to assign a Gibbs free energy to each promoter state. For our system, it corresponds to a 4-vector \mathbf{G}_0 in the standard condition (i.e., assume that all TFs have unit concentration. For arbitrary concentrations, $\mathbf{G} = \mathbf{G}_0 + k \sum_{f \in S} \ln [f]$, where $[f]$ represents the TF concentration, S the set of TFs that are bound to the promoter at a given moment, and k is a constant related to the Boltzmann factor. Note that our

model does not consider the second term in the total G since it does not consider TF regulations). This representation allows one to predict equilibrium steady-states (by applying the Boltzmann factor) and has been widely used to investigate the mean aspects of prokaryotic regulation^{36,81}. But it has the drawback to restrict the analysis to energetically-closed systems and, not carrying any kinetic information, it forbids any investigation of the stochastic aspects of gene expression. For this energetic formulation to be equivalent to the kinetic one, one has to consider an additional set of energy values, which however are difficult to access experimentally, namely the energy of the activation barrier for each reaction.

References

- Ptashne, M. Gene regulation by proteins acting nearby and at a distance. *Nature* **322**, 697–701 (1986).
- Whalen, S., Truty, R. M. & Pollard, K. S. Enhancer-promoter interactions are encoded by complex genomic signatures on looping chromatin. *Nat. Genet.* **48**, 488–496 (2016).
- Grossman, S. R. *et al.* Systematic dissection of genomic features determining transcription factor binding and enhancer function. *Proc. Natl. Acad. Sci. USA* **114**, E1291–E1300 (2017).
- Dao, L. T. M. *et al.* Genome-wide characterization of mammalian promoters with distal enhancer functions. *Nat. Genet.* **49**, 1073–1081 (2017).
- Krivega, I. & Dean, A. Enhancer and promoter interactions-long distance calls. *Curr. Opin. Genet. Dev.* **22**, 79–85 (2012).
- Kwon, D. *et al.* Enhancer-promoter communication at the *Drosophila* engrailed locus. *Development* **136**, 3067–3075 (2009).
- Calhoun, V. C., Stathopoulos, A. & Levine, M. Promoter-proximal tethering elements regulate enhancer-promoter specificity in the *Drosophila* Antennapedia complex. *Proc. Natl. Acad. Sci. USA* **99**, 9243–9247 (2002).
- Cui, L., Murchland, I., Shearwin, K. E. & Dodd, I. B. Enhancer-like long-range transcriptional activation by λ CI-mediated DNA looping. *Proc. Natl. Acad. Sci. USA* **110**, 2922–2927 (2013).
- Li, Q., Barkess, G. & Qian, H. Chromatin looping and the probability of transcription. *Trends. In. Genet.* **22**, 197–202 (2006).
- Kim, T. K. & Shiekhattar, R. Architectural and functional commonalities between enhancers and promoters. *Cell* **162**, 948–959 (2015).
- Priesta, D. G. *et al.* Quantitation of interactions between two DNA loops demonstrates loop domain insulation in *E. coli* cells. *Proc. Natl. Acad. Sci. USA* **111**, 4449–4457 (2014).
- Hnisz, D., Day, D. S. & Young, R. A. Insulated neighborhoods: structural and functional units of mammalian gene control. *Cell* **167**, 1188–1200 (2016).
- Savitskaya, E. *et al.* Study of long-distance functional interactions between Su(Hw) insulators that can regulate enhancer-promoter communication in *Drosophila melanogaster*. *Mol. Cell. Biol.* **26**, 754–761 (2006).
- Doyle, B., Fudenberg, G., Imakaev, M. & Mirny, L. A. Chromatin loops as allosteric modulators of enhancer-promoter interactions. *PLoS Comput. Biol.* **10**, e1003867 (2014).
- Vernimmen, D. & Bickmore, W. A. The hierarchy of transcriptional activation: from enhancer to promoter. *Trends. In. Genet.* **31**, 696–708 (2015).
- Dowen, J. M. *et al.* Multiple structural maintenance of chromosome complexes at transcriptional regulatory elements. *Stem. Cell. Reports.* **1**, 371–378 (2013).
- Ong, C. T. & Corces, V. G. Enhancer function: new insights into the regulation of tissue-specific gene expression. *Nat. Rev. Genet.* **12**, 283–293 (2011).
- Liu, T., Zhang, J. & Zhou, T. Effect of interaction between chromatin loops on cell-to-cell variability in gene expression. *PLoS Comput. Biol.* **12**, e1004917 (2016).
- Sánchez, A., Garcia, H. G., Jones, D., Phillips, R. & Kondev, J. Effect of promoter architecture on the cell-to-cell variability in gene expression. *PLoS Comput. Biol.* **7**, e1001100 (2011).
- Ddean, A. On a chromosome far, far away: LCRs and gene expression. *Trends. In. Genet.* **22**, 38–45 (2006).
- Batenchuk, C. *et al.* Chromosomal position effects are linked to sir2-mediated variation in transcriptional burst size. *Biophys. J.* **100**, 56–58 (2011).
- Blinka, S., Jr., R., M., Pulakanti, K. & Rao, S. Super-Enhancers at the Nanog Locus Differentially Regulate Neighboring Pluripotency-Associated Genes. *Cell Reports* **17**, 19 (2016).
- Bartman, C. R., Hsu, S. C., Hsiung, C. C., Raj, A. & Blobel, G. A. Enhancer regulation of transcriptional bursting parameters revealed by forced chromatin looping. *Mol. Cell.* **62**, 237–247 (2016).
- Wagner, A. Energy constraints on the evolution of gene expression. *Mol. Biol. Evol.* **22**, 1365–1374 (2005).
- Lease, R. A. & Belfort, M. A trans-acting RNA as a control switch in *Escherichia coli*: DsrA modulates function by forming alternative structures. *Proc. Natl. Acad. Sci. USA* **97**, 9919–9924 (2000).
- Li, C. & Wang, J. Landscape and flux reveal a new global view and physical quantification of mammalian cell cycle. *Proc. Natl. Acad. Sci. USA* **111**, 14130–14135 (2014).
- Schwanhäusser, B. *et al.* Global quantification of mammalian gene expression control. *Nature* **473**, 337–342 (2011).
- Brophy, J. A. & Voigt, C. A. Principles of genetic circuit design. *Nat. Methods* **11**, 508–520 (2014).
- Wong, J. J. L. *et al.* Orchestrated intron retention regulates normal granulocyte differentiation. *Cell* **154**, 583–595 (2013).
- Wang, J., Zhang, K., Xu, L. & Wang, E. Quantifying the waddington landscape and biological paths for development and differentiation. *Proc. Natl. Acad. Sci. USA* **108**, 8257–8262 (2011).
- Chen, Y. J., Johnson, S., Mulligan, P., Spakowitz, A. J. & Phillips, R. Modulation of DNA loop lifetimes by the free energy of loop formation. *Proc. Natl. Acad. Sci. USA* **111**, 17396–17401 (2015).
- Coulon, A., Chow, C. C., Singer, R. H. & Larson, D. R. Eukaryotic transcriptional dynamics: from single molecules to cell populations. *Nat. Rev. Genet.* **14**, 572–584 (2013).
- Vologodskii, A. V. & Cozzarelli, N. R. Conformational and thermodynamic properties of supercoiled DNA. *J. Mol. Biol.* **227**, 1224–1243 (1994).
- Mulligan, P., Chen, Y. J., Phillips, R. & Spakowitz, A. Interplay of protein binding interactions, DNA mechanics, and entropy in DNA looping kinetics. *Biophys. J.* **109**, 618–629 (2015).
- Boedicker, J. Q., Garcia, H. G. & Phillips, R. Theoretical and experimental dissection of DNA loop-mediated repression. *Phys. Rev. Lett.* **110**, 018101 (2013).
- Bintu, L. *et al.* Transcriptional regulation by the numbers: models. *Curr. Opin. Genet. Dev.* **15**, 125–135 (2005).
- Mehta, P. & Schwab, D. J. Energetic costs of cellular computation. *Proc. Natl. Acad. Sci. USA* **109**, 17978–17982 (2012).
- Zhang, X. J., Qian, H. & Qian, M. Stochastic theory of nonequilibrium steady states and its applications part I. *Phys. Rep.* **510**, 1–86 (2012).
- Qian, H. Open-system nonequilibrium steady state: statistical thermodynamics, fluctuations, and chemical oscillations. *J. Phys. Chem. B.* **110**, 15063–15074 (2006).
- Shahrezaei, V. & Swain, P. S. Analytical distributions for stochastic gene expression. *Proc. Natl. Acad. Sci. USA* **105**, 17256–17261 (2008).
- Paulsson, J. Models of stochastic gene expression. *Phys. Life Rev.* **2**, 157–175 (2005).

42. Zhang, J. J. & Zhou, T. S. Promoter-mediated transcriptional dynamics. *Biophys J.* **106**, 479–488 (2014).
43. Gillespie, D. T. Exact Stochastic Simulation of Coupled Chemical Reactions. *J. Chem. Phys.* **81**, 2340–2361 (1977).
44. Gillespie, D. T. Approximate accelerated stochastic simulation of chemically reacting systems. *J. Chem. Phys.* **115**, 1716–1733 (2001).
45. Kundu, S. & Peterson, C. L. Role of chromatin states in transcriptional memory. *BBA-Gen. Subjects* **1790**, 445–455 (2009).
46. Hardie, D. G. AMP-activated protein kinase: an energy sensor that regulates all aspects of cell function. *Genes. Dev.* **25**, 1895–908 (2011).
47. Cao, Y., Wang, H., Ouyang, Q. & Tu, Y. The free-energy cost of accurate biochemical oscillations. *Nat. Phys.* **11**, 772–778 (2015).
48. Schulze, J. O. *et al.* Bidirectional allosteric communication between the ATP-binding site and the regulatory PIF pocket in PDK1 protein kinase. *Cell. Chem. Biol.* **23**, 1193–1205 (2016).
49. Lebowitz, J. & Spohn, H. A Gallavotti-Cohen-type symmetry in the large deviation functional for stochastic dynamics. *J. Stat. Phys.* **95**, 333–365 (1999).
50. Lan, G., Sartori, P., Neumann, S., Sourjik, V. & Tu, Y. The energy-speed accuracy trade-off in sensory adaptation. *Nat. Phys.* **8**, 422–428 (2012).
51. Sartori, P. & Tu, Y. Free energy cost of reducing noise while maintaining a high sensitivity. *Phys. Rev. Lett.* **115**, 118102 (2015).
52. Qian, H. Phosphorylation energy hypothesis: open chemical systems and their biological functions. *Annu. Rev. Phys. Chem.* **58**, 113–142 (2007).
53. Ge, H. & Qian, H. Physical origins of entropy production, free energy dissipation, and their mathematical representations. *Phys. Rev. E* **81**, 561–578 (2010).
54. Tuan, D. *et al.* Transcription of the hypersensitive site HS2 enhancer in erythroid cells. *Proc. Natl. Acad. Sci. USA* **89**, 11219–11223 (1992).
55. Rippe, K. *et al.* Action at a distance: DNA-looping and initiation of transcription. *Trends Biochem. Sci.* **20**, 500–506 (1995).
56. Dekker, J. & Heard, E. Structural and functional diversity of topologically associating domains. *FEBS letters* **589**, 2877–2884 (2015).
57. Smemo, S. *et al.* Obesity-associated variants within FTO form long-range functional connections with IRX3. *Nature* **507**, 371–375 (2014).
58. Fukaya, T., Lim, B. & Levine, M. Enhancer Control of Transcriptional Bursting. *Cell* **166**, 358 (2016).
59. Ahsendorf, T., Wong, F., Eils, R. & Gunawardena, J. A framework for modeling gene regulation which accommodates non-equilibrium mechanisms. *BMC. Biol.* **12**, 1–23 (2014).
60. Gulyaev, A. P., van Batenburg, F. H. & Pleij, C. W. An approximation of loop free energy values of RNA H-pseudoknots. *RNA* **5**, 609–617 (1999).
61. Kieffer-Kwon, K. R. *et al.* Interactome maps of mouse gene regulatory domains reveal basic principles of transcriptional regulation. *Cell* **155**, 1507–1520 (2013).
62. Shen, Y. *et al.* A map of the cis-regulatory sequences in the mouse genome. *Nature* **488**, 116–120 (2012).
63. Nord, A. S. *et al.* Rapid and pervasive changes in genome-wide enhancer usage during mammalian development. *Cell* **155**, 1521–1531 (2011).
64. Ho, Y., Elefant, F., Liebhaber, S. A. & Cooke, N. E. Locus control region transcription plays an active role in long-range gene activation. *Mol. Cell.* **23**, 365–375 (2006).
65. Zhao, H., Friedman, R. D. & Fournier, R. E. The locus control region activates serpin gene expression through recruitment of liver-specific transcription factors and RNA polymerase II. *Mol. Cell. Biol.* **27**, 5286–5295 (2007).
66. Ghavi-Helm, Y. *et al.* Enhancer loops appear stable during development and are associated with paused polymerase. *Nature* **512**, 96–100 (2014).
67. Szutorisz, H., Georgiou, A., Tora, L. & Dillon, N. The proteasome restricts permissive transcription at tissue-specific gene loci in embryonic stem cells. *Cell* **127**, 1375–1388 (2006).
68. Vernimmen, D. *et al.* Chromosome looping at the human alpha-globin locus is mediated via the major upstream regulatory element (HS-40). *Blood* **114**, 4253–4260 (2009).
69. Song, S. H. *et al.* Multiple functions of Ldb1 required for beta-globin activation during erythroid differentiation. *Blood* **116**, 2356–2364 (2010).
70. Lin, C., Garruss, A. S., Luo, Z., Guo, F. & Shilatifard, A. The RNA Pol II elongation factor Ell3 marks enhancers in ES cells and primes future gene activation. *Cell* **152**, 144–156 (2012).
71. Seenundun, S. *et al.* UTX mediates demethylation of H3K27me3 at muscle-specific genes during myogenesis. *EMBO. J.* **29**, 1401–1411 (2010).
72. Vernimmen, D. *et al.* Polycomb eviction as a new distant enhancer function. *Genes. Dev.* **25**, 1583–1588 (2011).
73. Taberlay, P. C. *et al.* Polycomb-repressed genes have permissive enhancers that initiate reprogramming. *Cell* **147**, 1283–1294 (2011).
74. Kondo, T. *et al.* Polycomb potentiates meis2 activation in midbrain by mediating interaction of the promoter with a tissue-specific enhancer. *Dev. Cell.* **28**, 94–101 (2014).
75. Vernimmen, D. Uncovering enhancer functions using the alpha-globin locus. *PLoS. Genet.* **10**, e1004668 (2014).
76. Lettice, L. A. *et al.* Disruption of a long-range cis-acting regulator for Shh causes preaxial polydactyly. *Proc. Natl. Acad. Sci. USA* **99**, 7548–7553 (2002).
77. Nobrega, M. A., Ovcharenko, I., Afzal, V. & Rubin, E. M. Scanning human gene deserts for long-range enhancers. *Science* **302**, 413 (2003).
78. Jin, F. *et al.* A high-resolution map of the three-dimensional chromatin interactome in human cells. *Nature* **503**, 290–294 (2013).
79. Maeda, R. K. & Karch, F. Ensuring enhancer fidelity: Scanning human gene deserts for long-range enhancers. *Nat. Genet.* **34**, 360–361 (2003).
80. Li, G. *et al.* Extensive promoter-centered chromatin interactions provide a topological basis for transcription regulation. *Cell* **148**, 84–98 (2012).
81. Ackers, G., Johnson, A. & Shea, M. Quantitative model for gene regulation by lambda phage repressor. *Proc. Natl. Acad. Sci. USA* **79**, 1129–1133 (1982).

Acknowledgements

This work was supported by Grant Nos 11701117(LF), 2017A030310590(LF), 11631005(JS), 11626246(JS), 11471085(JS), 91530320(TS), 91230204(TS), 2014CB964703(TS), 11501193(LF) and 11761025(LF), IRT_16R16(JS), P. R. China.

Author Contributions

Conceived and designed the experiments: L.H., T.Z., J.Y. Performed the experiments: L.H. Analyzed the data: L.H., P.L., Z.Y. Contributed reagents/materials/analysis tools: L.H., J.Y., T.Z. Wrote the paper: T.Z.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-017-12765-x>.

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017