Supplementary Information: Transcription rates in DNA brushes

Tetsuya Yamamoto and S. A. Safran

DOI: 10.1039/c4sm02871f

1 Diffusion of RNA polymerase

For the cases that transcription is suppressed, the local concentrations of RNA polymerase (RNAP) in the solution of the brush region are derived by minimizing the free energy that has the form

$$\frac{F_{\rm bru}}{T} = \int dz \left[(\rho(z) \log \rho(z) - \rho(z)) + v\Phi(z)\rho(z) \right]. \tag{1}$$

The first term is free energy contributions due to the translational entropy of RNAP and the second term is free energy contributions due to the interactions between DNA chain segments and RNAP. We here treat the case in which the concentrations of RNAP are relatively small in the solution of the brush region; the interactions between RNAP molecules are very small and the structures of the DNA brush are only negligibly modified by RNAP. The contributions of the interactions between RNAP and DNA chain segments to the structures of DNA brushes are shown in sec. 4 in this Supplementary Material. Eq. (S1) should be minimized with respect to the local concentrations $\rho(z)$ of RNAP. $\Phi(z)$ is the local concentrations of DNA chain segments. T is the absolute temperature in the unit of the Boltzmann constant. v is the second virial coefficient that accounts for the interactions between RNAP and DNA chain segments.

The (local) chemical potentials $\mu(z)$ ($\equiv \delta F_{\rm bru}/\delta\rho(z)$) of RNAP thus have the form

$$\frac{\mu(z)}{T} = \log \rho(z) + v\Phi(z). \tag{2}$$

The flux arising from the gradient of chemical potentials have the form

$$J = -D\rho(z)\frac{\partial}{\partial z}\left(\frac{\mu(z)}{T}\right), \tag{3}$$

where D is the diffusion constant of RNAP in the solution of the brush region (eq. (S3) is derived by using the Einstein relationship). Substituting eq. (S2) into eq. (S3) leads to the form

$$J = -D\left[\frac{\partial}{\partial z}\rho(z) + v\rho(z)\frac{\partial}{\partial z}\Phi(z)\right]. \tag{4}$$

The chemical potentials of RNAP in the solution above the brush region have the form $\mu_0 = T\log\rho_0$ because the concentrations of RNAP in the solution is small and the interactions between RNAP molecules are negligible. Because the chemical potentials of RNAP are continuous at the interface between the brush region and the solution above the brush, $\mu(h) = \mu_0$, the local concentrations of RNAP at this interface have the form

$$\rho(h) = \rho_0 {\rm e}^{-v\Phi(h)}. \tag{5}$$

This is one of the boundary conditions that is applied to the local concentrations of RNAP at $z = h$.

2 Formal derivation of transcription dipoles

We treat DNA that has the promoter and terminator sites at s_p-th and s_t-th chain segments (here and after, the subscripts and superscripts "p" and "t" indicate the promoters and terminators, respectively). We use a positional vector $\mathbf{r}_\alpha(s) = (x_\alpha(s), y_\alpha(s), z_\alpha(s))$ to treat the conformation of the α-th DNA in a brush ($\alpha = 1, 2, \cdots, n$). Because the DNA brush is uniform in the lateral direction (the x-y plane), we treat only the z-component of these positional vectors. The local concentrations $\rho(z,t)$ of RNAP in the solution of the brush region are determined by a rate equation that has the form

$$\frac{\partial}{\partial t}\rho(z,t) = -\hat{g}_p(z)k^p_{\rm on}\rho(z,t) + k^p_{\rm off}\hat{g}_p(z)\hat{c}_p(z) + k^t_{\rm off}\hat{g}_t(z)\hat{c}_t(z). \tag{6}$$

$\hat{g}_p(z)$ and $\hat{g}_t(z)$ are the local concentration of promoters and terminators at z and have the forms

$$\hat{g}_p(z) = \sum_{\alpha=1}^{n} \delta(z - z_\alpha(s_p)) \tag{7}$$

$$\hat{g}_t(z) = \sum_{\alpha=1}^{n} \delta(z - z_\alpha(s_t)), \tag{8}$$

where n is the number of DNA chains in the brush. $\hat{c}_p(z)$ and $\hat{c}_t(z)$ are the fraction of promoters and terminators, to which RNAP is bound (and are located at z), and have the forms

$$\hat{c}_p(z) = \frac{1}{\hat{g}_p(z)} \sum_{\alpha=1}^{n} C_{p\alpha}\delta(z - z_\alpha(s_p)) \tag{9}$$

$$\hat{c}_t(z) = \frac{1}{\hat{g}_t(z)} \sum_{\alpha=1}^{n} C_{t\alpha}\delta(z - z_\alpha(s_t)). \tag{10}$$

$C_{p\alpha}(t)$ $(= C_\alpha(s_p,t))$ and $C_{t\alpha}(t)$ $(= C_\alpha(s_t,t))$ are the (average) number of RNAP at the promoter and the terminator of α-th DNA chain and satisfy rate equations that have the form

$$\frac{\partial}{\partial t}C_\alpha(s_p,t) = k^p_{\rm on}\rho(z_\alpha(s_p),t) - k^p_{\rm off}C_\alpha(s_p,t) - \xi C_\alpha(s_p,t) \tag{11}$$

$$\frac{\partial}{\partial t}C_\alpha(s,t) = -\xi C_\alpha(s,t) + \xi C_\alpha(s-\Delta a,t) \tag{12}$$

$$\frac{\partial}{\partial t}C_\alpha(s_t,t) = -k^t_{\rm off}C_\alpha(s_t,t) + \xi C_\alpha(s_t-\Delta a,t), \tag{13}$$

where eqs. (S11), (S12), and (S13) are effective for $s = s_p$, $s_p < s < s_t$, and $s = s_t$, respectively. $C_\alpha(s,t)$ is the number of RNAP at the s-th chain segment of the α-th DNA chain (at time t). Δa is the distance between the two neighbouring bases.

We solve eqs. (S6), (S11), (S12), and (S13) for stationary states. For the case in which RNAP is released from the terminators slowly ($k^t_{\rm off}$ is very small), RNAP molecules may be stacked in TX units. In these cases, the excluded volume interactions between RNAP in a TX unit may limit the dynamics of transcription. We here treat a simple case in which RNAP is released from terminators relatively fast and the excluded volume interactions between RNAP molecules thus do not limit the dynamics of transcription. In steady states, eqs. (S11), (S12), and (S13) lead to the relationships

$$k^t_{\rm off}C_{t\alpha} = \lambda\rho(z_\alpha(s_p)) \tag{14}$$

$$k^p_{\rm on}\rho(z_\alpha(s_p)) - k^p_{\rm off}C_{p\alpha} = \lambda\rho(z_\alpha(s_p)) \tag{15}$$

with

$$\lambda \equiv \frac{\xi}{k_{\rm off}^p + \xi} k_{\rm on}^p. \tag{16}$$

This leads to the form

$$\begin{aligned} -\hat{g}_p(z) k_{\rm on}^p \rho(z,t) + k_{\rm off}^p \hat{g}_p(z) \hat{c}_p(z) &= -\sum_{\alpha=1}^{n} (k_{\rm on}^p \rho(z,t) - k_{\rm off}^p C_{p\alpha}) \delta(z - z_\alpha(s_p)) \\ &= -\lambda \sum_{\alpha=1}^{n} \rho(z) \delta(z - z_\alpha(s_p)) \\ &= -\lambda \sum_{\alpha=1}^{n} \rho(z_\alpha(s_p)) \delta(z - z_\alpha(s_p)) \qquad (17) \\ k_{\rm off}^t \hat{g}_t(z) \hat{c}_t(z) &= \lambda \sum_{\alpha=1}^{n} \rho(z_\alpha(s_p)) \delta(z - z_\alpha(s_t)). \qquad (18) \end{aligned}$$

For the case in which the length $l_{\rm TX}$ of TX units (equivalently, the separation between the promoter and terminator sites of TX units) is smaller than the Kuhn length $l_{\rm a}$ of DNA chains, the positions of the promoters and terminators of TX units have an approximate relationship that has the form

$$z_\alpha(s_t) = z_\alpha(s_p) + u_{z\alpha} l_{\rm TX}, \tag{19}$$

where $u_{z\alpha}$ is the z-component of tangent vector of DNA chains at the positions of their TX units (we here note that this tangent vector only depends on the index α of DNA chains). Substituting eqs. (S17) and (S18) into eq. (S6) leads to the form

$$\begin{aligned} \frac{\partial}{\partial t} \rho(z,t) &= -\lambda \sum_{\alpha=1}^{n} \rho(z_\alpha(s_p)) \delta(z - z_\alpha(s_p)) + \lambda \sum_{\alpha=1}^{n} \rho(z_\alpha(s_p)) \delta(z - z_\alpha(s_t)) \\ &= -\lambda \sum_{\alpha=1}^{n} \rho(z_\alpha(s_p)) \left[\delta(z - z_\alpha(s_p)) - \delta(z - z_\alpha(s_t)) \right] \\ &\simeq -\frac{\partial}{\partial z} \hat{P}_z \qquad (20) \end{aligned}$$

with

$$\hat{P}_z(z) = l_{\rm TX} \lambda \sum_{\alpha=1}^{n} u_{z\alpha} \rho(z_\alpha(s_p)) \delta(z - z_\alpha(s_p)) \tag{21}$$

where we used eq. (S19) to derive the last form of (S20). $\hat{P}_z(z)$ is the density of TX dipoles and is rewritten in the form

$$P_z(z) = l_{\rm TX} \lambda S_1(z) \rho(z) g_p(z), \tag{22}$$

where $S_1(z)$ is the local average of the z-component of the tangent vector of DNA that has its TX unit at a position z and has the form

$$S_1(z) = \frac{1}{g_p(z)} \sum_{\alpha=1}^{n} u_{z\alpha} \delta(z - z_\alpha(s_p)). \tag{23}$$

3 Uniform brush

3.1 The orientational order parameter

We here treat a simple case in which the concentrations of DNA chain segments are uniform in the brush region. This DNA brush is treated by using a free energy (per unit area) that has the form

$$\begin{aligned}\frac{F}{\sigma T} &= N\int d\Omega n(\mathbf{u})\log n(\mathbf{u}) \\ &\quad +\frac{1}{2}N\Phi_0\int d\Omega_i\int d\Omega_j\beta_{ij}n(\mathbf{u}_i)n(\mathbf{u}_j).\end{aligned} \tag{24}$$

The first term is a free energy contribution due to the orientational entropy of DNA chain segments and the second term is a free energy contribution due to the anisotropic excluded volume interactions between DNA chain segments; this free energy takes into account the semiflexibility of DNA in an extension of the free energy of flexible polymer brush (see also eq. (S28)). $n(\mathbf{u})$ is the orientational distribution functions of DNA chain segments and $\mathbf{u}=(\sin\theta\cos\phi,\sin\theta\sin\phi,\cos\theta)$ is the unit vector that is parallel to DNA chain segments (θ is the angle between DNA chain segments and the normal of the substrate and ϕ is the azimuthal angle). The integral $d\Omega$ ($\equiv\sin\theta d\theta d\phi$) should be performed for all possible orientations $\mathbf{u}$ of DNA chain segments (the subscripts i and j represent the integrals for two interacting chain segments). This accounts for the chain segments in the bulk of the brush and thus is effective for the case in which the number N of chain segments per chain is relatively large. In this model, the substrate plays a role in breaking the symmetry of the system; although eq. (24) is symmetric with respect to the outward and inward normal to the substrate, we choose the solution, where DNA chains are stretched towards the outward normal to the substrate. N is the number of chain segments of each DNA. Φ_0 ($\equiv N\sigma/h$) is the concentrations of DNA chain segments in the DNA brush, where σ is the grafting density of DNA and h ($=Nl_{\rm a}S_1$) is the height of the brush ($l_{\rm a}$ is the Kuhn length of DNA). S_1 is the (first order) orientational order parameter that is defined by

$$S_1=\int d\Omega\cos\theta n(\mathbf{u}). \tag{25}$$

β_{ij} ($\equiv 2dl_{\rm a}^2|\mathbf{u}_i\times\mathbf{u}_j|$) is the second virial coefficient.

The orientational distribution function $n(\mathbf{u})$ is expanded in a series of Legendre polynomials in the form

$$n(\mathbf{u})=\frac{1}{2\pi}\sum_{k=1}^{\infty}\frac{2k+1}{2}S_kP_k(\cos\theta), \tag{26}$$

where $P_k(\cos\theta)$ ($k=0,1,2,\cdots$) is the Legendre polynomials of the k-th order and the coefficients

$$S_k=\int d\Omega P_k(\cos\theta)n(\mathbf{u}) \tag{27}$$

are the k-th order orientational order parameter. For the cases that the orientational order parameter S_1 is not very close to unity and there are no higher order orientational order parameters, eq. (S24) has an asymptotic form

$$\frac{F_{\rm nem}}{\sigma T}\simeq\frac{3}{2}\frac{h^2}{Nl_{\rm a}^2}+\frac{1}{2}w\frac{\sigma N^2}{h}, \tag{28}$$

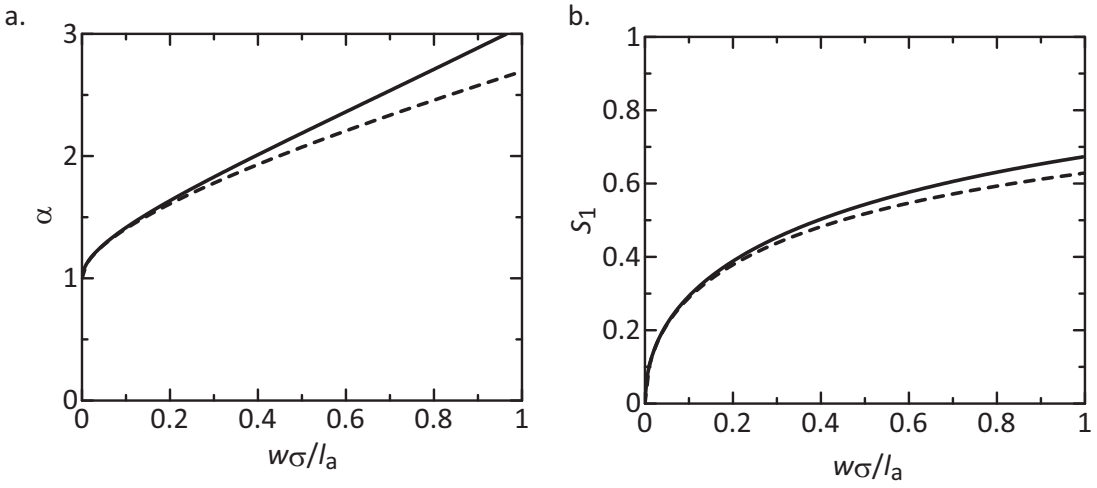

Fig. 1 The parameter α in eq. (S29) (a) and the first order orientational order parameter S_1 (b) are calculated as functions of the (rescaled) grafting density $w\sigma/l_a$ (solid). We used an asymptotic expression, eq.(S32) (for large grafting densities) to derive the broken curves.

where we used the second virial coefficients w $(=\pi d l_{TX}^2/2)$ for the cases that the orientational distributions of DNA chain segments are approximately isotropic. Eq. (S24) thus takes into account the anisotropic excluded volume interactions between DNA chain segments and the inextensibility of semiflexible polymers in an extension of a free energy of flexible polymer brush.

We use a trial function that has the form

$$n(\mathbf{u}) = \frac{\alpha}{2\pi} \mathrm{e}^{-\alpha\theta^2/2} \tag{29}$$

to derive an approximate form of the orientational distribution function $n(\mathbf{u})$, where α is a parameter that characterizes the orientational distribution of chain segments (a similar trial function was used by Onsager[1,2]). The Gaussian form of eq. (S29) implies that eq. (S29) is effective for a relatively large grafting density (where the deviations of the orientations DNA chain segments from the normal of the substrate are small). With this trial function, the first order orientational order parameter S_1 has an approximate form

$$S_1 \simeq 1 - \frac{1}{\alpha}, \tag{30}$$

see eq. (S25). The free energy thus has an approximate form

$$\frac{F_{\mathrm{nem}}}{N\sigma T} \simeq \log\alpha - 1 + \frac{3}{\sqrt{2\pi}} \frac{w\sigma}{l_a} \frac{\sqrt{\alpha}}{\alpha - 1}. \tag{31}$$

The parameter α is derived by minimizing the free energy, eq. (S31) with respect to α, see fig. S1. For small (rescaled) grafting densities, the parameter α has an asymptotic form

$$\alpha = 1 + \sqrt{\frac{3}{\sqrt{2\pi}}} \sqrt{\frac{w\sigma}{l_a}} + \frac{1}{2} \frac{3}{\sqrt{2\pi}} \frac{w\sigma}{l_a}, \tag{32}$$

see the broken curves in fig. S1.

3.2 The local concentrations of RNAP in the brush and transcription rate

We here treat the case in which each DNA has one TX unit at its s_0-th chain segment. For the cases that the concentrations of DNA chain segments are uniform in the brush region, Φ_0 $(=N\sigma/h)$, the local density of RNA polymerases in the form

$$\rho(z) = \rho_0 \exp\left(-\nu\Phi_0 - \frac{\lambda l_{\rm TX} S_1}{D}\int_z^h dz' g_{s_0}(z')\right), \tag{33}$$

where we used eq. (3) in the main article (and the boundary condition for the local concentrations of RNAP, eq. (S5)). We here continue our calculation without specifying the form of the local concentrations $g_{s_0}(z)$ of TX units (see also eq. (S36) below). Eq. (S33) predicts that the local concentrations of RNAP depend on the position s_0 of TX units along DNA. Substituting eq. (S33) into eq. (5) in the main article leads to TX rates in the form

$$\begin{aligned} R &= \frac{D}{l_{\rm TX}S_1\sigma}\rho_0 {\rm e}^{-\nu\Phi_0}\int_0^h dz \frac{\partial}{\partial z}\left[\exp\left(-\frac{\lambda l_{\rm TX}S_1}{D}\int_z^h dz' g_{s_0}(z')\right)\right] \\ &= \frac{D}{l_{\rm TX}S_1}\rho_0 {\rm e}^{-\nu\Phi_0}\left(1-{\rm e}^{-\lambda l_{\rm TX}S_1\sigma/D}\right). \end{aligned} \tag{34}$$

The last form of eq. (S34) is eq. (6) in the main article. Indeed, this is a general result that does not depend on the specific form of the local concentrations $g_{s_0}(z)$ of promoters; this results from only the fact that each DNA has one promoter

$$\int_0^h dz g_{s_0}(z) = \sigma. \tag{35}$$

Our theory thus predicts that TX rates do not depend on the position s_0 of the promoters for the cases that the concentrations of DNA chain segments are uniform in the brush, see eq. (S34).

3.3 Fluctuations of the orientations of DNA chain segments: First approximation

We derive the form of the local concentrations $g_{s_0}(z)$ of the promoters by using the fact that the orientational distributions of DNA chain segments of uniform DNA brushes do not depend on the distance z from the substrate, see eq. (29). The central limit theorem predicts that the local (number) density of TX dipoles have the form

$$\begin{aligned} g_{s_0}(z) &= \langle \delta\left(z - l_{\rm a}\sum_{k=1}^{s_0}\cos\theta_k\right)\rangle \\ &= \frac{\sigma\alpha}{\sqrt{2\pi s_0 l_{\rm a}^2}}{\rm e}^{-\alpha^2(z-s_0 l_{\rm a}S_1)^2/(2s_0 l_{\rm a}^2)}. \end{aligned} \tag{36}$$

Eq. (S36) predicts that the distribution of TX units becomes broader with moving the positions of the TX units from the grafted end to the free end of DNA. The central limit theorem is exact for large values of s_0 and thus eq. (36) is only an approximate form for small values of s_0 (however, we note that eq.(S36) returns to the delta function for $s_0 \to 0$).

The form of eq. (S36) is the distribution of the s_0-th chain segment of DNA in the brush; the local concentrations of chain segments is rederived by using eq. (S36) in the form

$$\begin{aligned}\Phi(z) &= \int_0^N ds \frac{\sigma\alpha}{\sqrt{2\pi s l_{\rm a}^2}} {\rm e}^{-\alpha^2(z-sl_{\rm a}S_1)^2/(2sl_{\rm a}^2)} \\ &\simeq \frac{\sigma}{2l_{\rm a}S_1}\left[1-{\rm Erf}\left(\frac{\alpha(z-Nl_{\rm a}S_1)}{\sqrt{2Nl_{\rm a}^2}}\right)\right], \end{aligned} \tag{37}$$

where we omitted a term

$$\frac{\sigma}{2l_{\rm a}S_1}{\rm e}^{2\alpha^2 zS_1/l_{\rm a}}\left[{\rm Erf}\left(\frac{\alpha(z+Nl_{\rm a}S_1)}{\sqrt{2Nl_{\rm a}^2}}\right)-1\right] \tag{38}$$

to derive the last form of eq. (S37). Eq. (S38) is small for the case in which the value of $\alpha S_1\sqrt{N}$ is large. Eq. (S37) returns to Φ_0 $(=\sigma/(l_{\rm a}S_1))$ for $z<h$; Φ_0 is indeed an asymptotic concentration of DNA chain segments in the bulk of the brush and these concentrations decrease to zero with the error function at $z\sim h$. Eq. (S37) is thus the first approximation that takes into account the finite distribution of the orientations of DNA chain segments. The fact that the local concentrations of DNA chain segments are smaller at the interface (between the brush region and the bulk of the solution), $z\sim h$, than in the bulk of the brush implies that the orientational order parameter S_1 is smaller at the interface than in the bulk. This is not taken into account our treatment that uses an orientational distribution function that does not depend on the distance z from the substrate. Our treatment is thus only an approximation for the case in which TX units is located at (the vicinity of) the free end of DNA.

4 Parabolic brush

For a moderate grafting density, the conformation of DNA in the brush does not greatly deviate the most probable conformations[3–5]. We thus use the classical approximation in which the fluctuations of DNA conformation from the most probable one are neglected[3–5]. Self-consistent field theories show that the local concentrations of DNA chain segments are quadratic function of the distance z from the substrate (parabolic brush)[3–5]. We here take into account the interactions between RNAP and DNA chain segments in an extension of the self-consistent field theories of parabolic brushes. Because the treatment of parabolic brush is well documented and our extension is straight forward, we do not repeat the derivation here (interested readers should refer to refs.[3] and[5]).

In self-consistent field theories, the interactions between DNA chain segments are treated by using molecular fields $w\Phi(z)$ (w is the second virial coefficient that accounts for the interactions between DNA chain segments and $\Phi(z)$ is the local concentrations of DNA chain segments). We take into account the interactions between RNAP and DNA chain segments in an extension of these molecular fields:

$$\omega(z) = w\Phi(z) + v\rho(z), \tag{39}$$

where v is the second virial coefficient that accounts for the interactions between RNAP and DNA chain segments and $\rho(z)$ is the local concentration of RNAP. The local concentrations of DNA chain segments have the form

$$\Phi(z) = \frac{v}{w}(\rho_0-\rho(z)) + \frac{3}{2}\frac{\sigma N}{h_{\rm flx}}\frac{z_m^2-z^2}{h_{\rm flx}^2}, \tag{40}$$

where σ is the grafting density of DNA and N is the number of chain segments of each DNA. ρ_0 is the concentration of RNAP in the bulk of the solution. $h_{\rm flx}$ is defined by the form

$$h_{\rm flx} = N l_{\rm a} \left(\frac{4}{\pi^2} \frac{w\sigma}{l_{\rm a}} \right)^{1/3} \tag{41}$$

z_m is the height of the brush for the case in which there are RNAP molecules and is determined by a relationship that has the form

$$\frac{z_m^3}{h_{\rm flx}^3} + \frac{v}{w\sigma N} \left(z_m \rho_0 - \int_0^{z_m} dz \rho(z) \right) = 1. \tag{42}$$

The height z_m returns to $h_{\rm flx}$ for the cases that there are no RNAP molecules in the system. The expressions of eq. (S42) ensure that the number of chain segments in the brush is conserved,

$$\int_0^{z_m} dz \Phi(z) = \sigma N. \tag{43}$$

The local concentrations of the free ends of DNA chains have the form

$$g_0(z) = 3\sigma \frac{z\sqrt{z_m^2 - z^2}}{h_{\rm flx}^3} + \frac{v}{wN} \int_z^{z_m} dz' \frac{z}{\sqrt{z'^2 - z^2}} \left(\frac{\partial}{\partial z'} \rho(z') \right). \tag{44}$$

Eq. (S42) ensures that the number of DNA chains is conserved;

$$\int_0^{z_m} dz g_0(z) = \sigma. \tag{45}$$

The most probable conformation of DNA that has free ends at $z = z_0$ is represented by using positional vectors that have the form $\mathbf{r}(s, z_0) = (x(s, z_0), y(s, z_0), z(s, z_0))$, where $z(s, z_0)$ has the form

$$z(s, z_0) = z_0 \sin\left(\frac{\pi s}{2N} \right). \tag{46}$$

$x(s, z_0)$ and $y(s, z_0)$ are not important because the brush is uniform in the lateral direction (the *x*-*y* plane). s is the index of chain segments that are count from the grafted ends (consistent with sec. 3). Indeed, the optimal conformations are not affected by interactions between DNA chain segments and RNAP because of the virtue of the classical approximation[5]. The tangent vectors of these conformations have the form $\mathbf{u}(s, z_0)$ $(\equiv (u_x(s, z_0), u_y(s, z_0), u_z(s, z_0)))$ with

$$u_z(s, z_0) = \frac{\pi}{2} \frac{z_0}{N l_{\rm a}} \cos\left(\frac{\pi s}{2N} \right). \tag{47}$$

The free energy of the brush (per unit area) has the form

$$\frac{F_{\rm bru}}{T} = F_0 + \frac{3}{2} v \int_0^{z_m} dz' \frac{N\sigma}{h_{\rm flx}} \frac{z_m^2 - z'^2}{h_{\rm flx}^2} \rho(z') - \frac{1}{2} \frac{v^2}{w} \int_0^{z_m} dz \rho^2(z), \tag{48}$$

where F_0 is free energy contributions

$$F_0 = -\frac{3\pi^2}{8} \frac{\sigma z_m^2}{N l_{\rm a}^2} \left(\frac{2}{5} \frac{z_m^3}{h_{\rm flx}^3} - 1 \right) + \frac{1}{2} \frac{v^2}{w} z_m \rho_0^2 \tag{49}$$

that only indirectly depend on the total number of RNAP in the brush via eq. (S42). The third term of eq. (S48) is effective interactions between RNAP molecules and is very small for the cases that the concentrations of RNAP are small. Omitting these higher order terms with respect to the local concentrations $\rho(z)$ of RNAP leads to the form of the second term of eq. (S1) for parabolic brush. With the free energy contributions of the translational entropy of RNAP, see the first term of eq. (S1), eq. (S48) leads to the osmotic pressures in the form

$$\begin{aligned}\frac{\Pi_{\rm bru}}{T} &= -\rho_0\left[\log\rho_0-1\right] \\ &\quad +\frac{\pi^2}{2}\frac{3\sigma z_m}{2l_{\rm a}^2N}\left[\frac{z_m^3}{h_{\rm flx}^3}+\frac{v}{N\sigma w}\left(\rho_0 z_m-\int_0^{z_m}dz\rho(z)\right)-1\right].\end{aligned} \tag{50}$$

Eq. (S42) ensures the equality of osmotic pressures at the interface between the brush region and the solution.

We treat the cases that each DNA has one TX unit at its s_0-th segment. Eq. (S46) shows that DNA that has TX unit at a position z have their free ends at $z_0=\chi z$, where we used a parameter

$$\chi^{-1}=\sin\left(\frac{\pi s_0}{2N}\right) \tag{51}$$

to simplify the notation; because we only retain the most probable conformation for each free end positions in our calculations, the positions of TX units (the s_0-th chain segments) are uniquely determined by the positions of free ends. The local (number) concentration of TX units has the form

$$g_{s_0}(z)=3\sigma\frac{\chi^2 z\sqrt{z_m^2-\chi^2z^2}}{h_{\rm flx}^3}+\frac{v}{wN}\int_{\chi z}^{z_m}dz'\frac{\chi^2 z}{\sqrt{z'^2-\chi^2z^2}}\left(\frac{\partial}{\partial z'}\rho(z')\right) \tag{52}$$

for $0<z<\chi^{-1}z_m$ and $g_{s_0}(z)=0$ for $\chi^{-1}z_m<z<z_m$. The orientation of TX units of DNA is parallel (the 'OUT' configuration) or anti-parallel (the 'IN' configuration) to the tangent vector of the DNA, eq. (S47). TX units that are located at z has the local projection along the z direction that has the form

$$u_z(z)=\frac{\pi}{2}\frac{z}{Nl_{\rm a}}\cot\left(\frac{\pi s_0}{2N}\right). \tag{53}$$

Because we only retain the most probable conformation of DNA for each free end positions in our calculations, all of TX units at the same position z show the same orientation; $S_1(z)=u_z(z)$.

5 Brushes of non-interacting hard-rods

We here analyze a DNA brush, where each DNA chain is treated as a hard-rod (in contrast to the approximation of flexible chain) and the concentration of these rods is relatively dilute and the interactions between different rods are negligible. In this case, the orientations of the DNA rods are limited to the space that is above the substrate (where the angle θ between a rod and the outward normal to the substrate is in the range of $0<\theta<\pi/2$) and the distribution of the orientations is uniform. The orientational distribution function $n(\mathbf{u})$ thus has the form

$$n(\mathbf{u})=\frac{1}{2\pi}\Theta(\cos\theta), \tag{54}$$

where $\Theta(x)$ is 1 for $x > 0$ and 0 for $x < 0$. The local concentration of DNA chain "segments" has the form

$$\begin{aligned}
\Phi(z) &= \sigma \int_0^N ds\, \langle \delta\left(z - s l_{\rm a} \cos\theta\right) \rangle \\
&= \frac{\sigma}{l_{\rm a}} \left\langle \int_0^{L\cos\theta} du\, \frac{\delta(u-z)}{\cos\theta} \right\rangle \\
&= \frac{\sigma}{l_{\rm a}} \int_0^{\cos^{-1}(z/L)} d\theta \frac{\sin\theta}{\cos\theta} \\
&= -\frac{\sigma}{l_{\rm a}} \log\left(\frac{z}{L}\right),
\end{aligned} \tag{55}$$

where σ is the grafting density of DNA chains, L is the overall length of the rods, $\langle\rangle$ is the thermodynamic average with respect to the distribution function $n(\mathbf{u})$, and $\delta(x)$ is the Dirac delta function. The Kuhn length $l_{\rm a}$ of DNA is twice the persistence length of DNA, where this length is determined by the absolute temperature T and the bending rigidity of the rods, and the effective number N of chain segments is defined by $L/l_{\rm a}$ that can be less than unity; this treatment makes our model of hard-rods compatible with the formalism that is presented in the main article and other sections of this Supplementary Information. The local density of promoter sites has the form

$$\begin{aligned}
g_{s_0}(z) &= \langle \delta(z - s_0 l_{\rm a} \cos\theta) \rangle \\
&= \frac{\sigma}{l_{\rm a} s_0}
\end{aligned} \tag{56}$$

for $z < l_{\rm a} s_0$ and zero for $z > l_{\rm a} s_0$, for the cases in which each DNA chain has a TX unit at the s_0-th chain segment. The orientational order parameter $S_1(z)$ has the form

$$\begin{aligned}
S_1(z) &= \frac{1}{g_{s_0}(z)} \sigma \langle \cos\theta\ \delta(z - s_0 l_{\rm a} \cos\theta) \rangle \\
&= \frac{\sigma}{g_{s_0}(z)} \int_0^{\pi/2} d\theta\ \sin\theta\cos\theta\ \delta(z - s_0 l_{\rm a} \cos\theta) \\
&= \frac{\sigma}{g_{s_0}(z)} \frac{1}{(s_0 l_{\rm a})^2} \int_0^{s_0 l_{\rm a}} du\, u\, \delta(z-u) \\
&= \frac{z}{s_0 l_{\rm a}}
\end{aligned} \tag{57}$$

for $z < s_0 l_{\rm a}$ and $S_1(z) = 0$ for $z > s_0 l_{\rm a}$, see also eq. (S23).

Substituting eqs. (S55), (S56), and (S57) in eq. (3) in the main text leads to the expression of the local concentration of RNAP in the form

$$\rho(z) = \rho_0 \left(\frac{z}{l_{\rm a} N}\right)^{\mu} \exp\left[-\frac{\lambda \sigma l_{\rm TX} l_{\rm a}}{2D} \left(1 - \frac{z^2}{l_{\rm a}^2 s_0^2}\right)\right] \tag{58}$$

for $z < N s_0$ and

$$\rho(z) = \rho_0 \left(\frac{z}{l_{\rm a} N}\right)^{\mu} \tag{59}$$

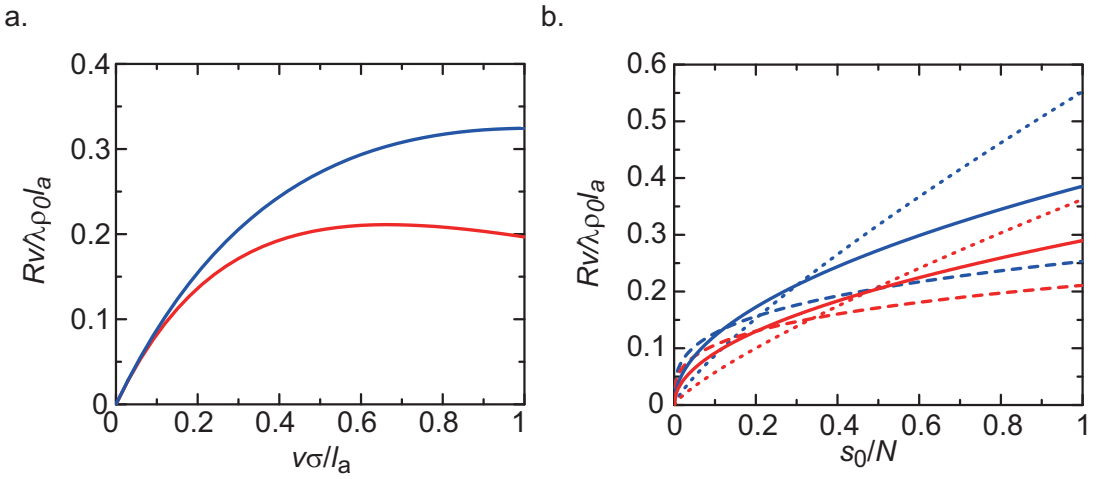

Fig. 2 a. The (rescaled) TX rate $Rv/(\lambda\rho_0 l_a)$ is calculated as a function of the (rescaled) grafting density $v\sigma/l_a$ by using eq. (S60) that is applicable to cases, in which each DNA behaves as hard-rods that are very dilutely packed so that interactions between these rods are negligible. The positions of TX units are fixed at $s_0 = 0.5$. The cases of both IN (blue) and OUT (red) orientations are shown. **b**. The (rescaled) TX rate $Rv/(\lambda\rho_0 l_a)$ is calculated as functions of the positions s_0/N of TX units along the DNA rod by using eq. (S60). The rescaled grafting densities $v\sigma/l_a$ that are used for the calculations are 0.3 (broken), 0.5 (solid), and 0.8 (dotted). The value of the (rescaled) rate constant $\lambda l_{TX} l_a/(Dv)$ is fixed at 1.0 for all the calculations.

for $Ns_0 < z < Nl_a$, where μ is the rescaled grafting density $v\sigma/l_a$. ρ_0 is the concentration of RNAP in the bulk of the solution and v is the second virial coefficient that accounts for DNA-RNAP interactions. D is the diffusion constant of RNAP in the solution in the brush region. λ is the rate constant for the binding of RNAP to a promoter site and l_{TX} is the length of a TX unit. The TX rate of the DNA brush has the form

$$\frac{R}{\lambda\sigma\rho_0} = \left(\frac{s_0}{N}\right)^{\mu} \int_0^1 du\, u^{\mu} \exp\left[\frac{\lambda\sigma l_{TX} l_a}{2D}(u^2-1)\right], \tag{60}$$

where this equation is derived by substituting eqs. (S58) and (S59) in eq. (5) in the main text. Eqs. (S58) and (S60) are applicable to the OUT orientation and the corresponding equations for the IN orientation are derived by the transformation $\lambda \to -\lambda$. Eq. (S60) predicts that the TX rate scales with the position s_0 of the TX units as $R \sim s_0^{\mu}$ and is thus only weakly dependent on the position of the TX units, s_0, for $\mu < 1$, which applies if the interactions between DNA chain segments and RNAP are small, see the broken curves in fig. 2 **b**; this weak dependence arises from the fact that the local concentration of DNA chain segments is only logarithmically dependent on the position z, see eq. (S55). This calculation also predicts that the TX rate *increases* as the TX units are shifted from the grafting ends to the free ends for both of the 'IN' and 'OUT' orientations; the latter is not the case of experiments of ref.[9]. In conclusion, brushes of non-interacting hard-rods are relatively uniform, where the concentration varies only logarithmically with the position z, see eq. (S55). However, this does not account for the experimental results of ref. [9] that show that for the IN orientation, the TX rate *decreases* as the TX units are shifted from the grafting ends to the free ends. A subtle balance between the flexibility and bending rigidity may be involved in the experiments that suggest that the TX rate of the DNA brush (that has been measured in ref.[9]) is not sensitive to the positions of TX units. The treatment here is applicable to cases, in which the overall length of hard rods is larger than the size of RNAP; otherwise, one must use eq. (7) in the main article with $S_1 = 1/2$.

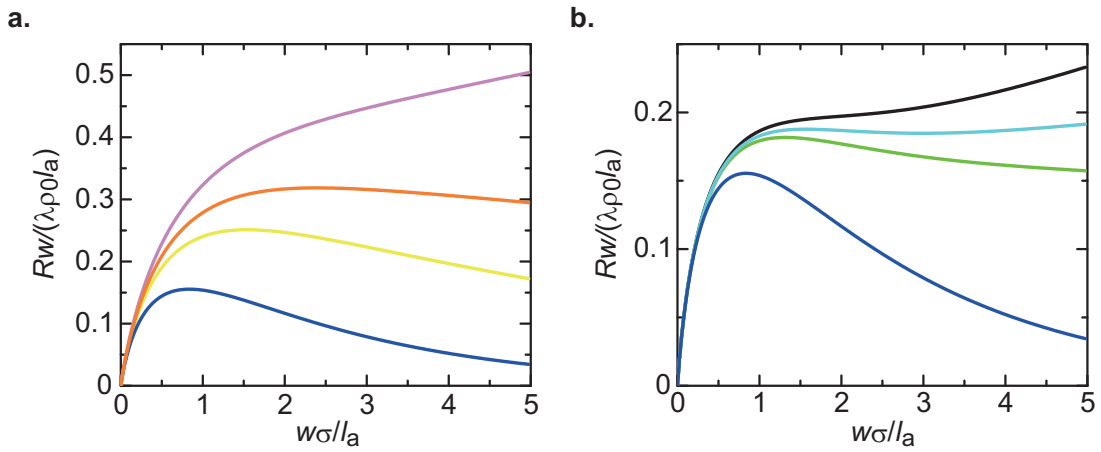

Fig. 3 a. The (rescaled) TX rate $Rw/(\lambda\rho_0 l_a)$ of a uniform DNA brush is shown as functions of the (rescaled) grafting density $w\sigma/l_a$ of DNA for the cases that the values of the ratio v/w of virial coefficients are 1.5 (blue), 1.2 (yellow), 1.1 (orange), and 1.0 (pink) in the IN configuration. The value of the (rescaled) rate constant $\lambda l_{TX} l_a/(Dw)$ is fixed to 1.0. **b**. The (rescaled) TX rate $Rw/(\lambda\rho_0 l_a)$ of a uniform DNA brush is shown as functions of the (rescaled) grafting density $w\sigma/l_a$ of DNA for the cases that the values of the (rescaled) rate constant $\lambda l_{TX} l_a/(Dw)$ are 1.0 (blue), 1.4 (light green), 1.45 (turquoise), and 1.5 (black) in the IN configuration. The value of the ratio v/w of the second virial coefficients is fixed to 1.5.

Symbol	Physical meaning	Order of magnitudes
v	2nd virial coefficient for DNA-RNAP interactions	3×10^4 nm^3
w	2nd virial coefficient for DNA-DNA interactions	3×10^4 nm^3
l_a	Kuhn length of DNA	100 nm
λ	Binding rate constant of RNAP to the promoters	1×10^{-19} s^{-1}m^{-3}
D	Diffusion coefficient of RNAP	3×10^{-11} m^2/s
l_{TX}	Length of TX units	95 nm

Table 1 The orders of magnitudes of parameters that are used in our theory are estimated for physiologically relevant salt concentration. We estimated the second virial coefficient v for DNA-RNAP interactions by treating RNAP as a hard sphere of radius 10 nm, where its volume is equal to the volume of RNAP of E-coli (an ellipsoid of 8.5 nm $\times$ 10.5 nm $\times$ 14 nm)[6]. With this treatment (and the viscosity of water 9×10^{-4} Pa s), the diffusion constant D of RNAP is estimated by using the Einstein's relationship. In general, the (mutual) diffusion constant is a function of the grafting density of DNA, but, for simplicity, we use the diffusion constant of the dilute limit. We used the values of the binding rate constant of RNAP to the promoters of TX units that was shown in ref.[7,8] and the values of the length of TX units that were used in experiments[9].

References

1 L. Onsager, *Ann. NY Acad. Sci.*, 1949, **51**, 627.
2 T. Odijk, *Macromolecules*, 1986, **19**, 2313.
3 R. R. Netz and M. Schick, *Macromolecules*, 1998, **31**, 5105.
4 S. T. Milner, T. A. Witten, and M. E. Cates, *Macromolecules*, 1988, **21**, 2610.
5 E. B. Zhulina, O. V. Borisov, and L. Brombacher, *Macromolecules*, 1991, **24**, 4679.
6 A. Polyakov, E. Severinova, and S. A. Darst, *Cell*, 1995, **83**, 365.
7 P. de Haseth, M. L. Zupancic, and M. Thomas Record Jr., *J. Bacteriol.*, 1998, **180**, 3019.
8 A. Újvári and C. T. Martin, 1996, *Biochemistry*, **35**, 14574.
9 S. S. Daube, D. Bracha, A. Buxboim, and R. H. Bar-Ziv, *Proc. Natl. Acad. Sci. USA*, 2010, **107**, 2836.