PHYSICAL REVIEW E **96**, 030402(R) (2017)

Osmotic mechanism of the loop extrusion process

Tetsuya Yamamoto*
Department of Materials Physics, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, 464-8603, Japan

Helmut Schiessel
Instituut-Lorentz for Theoretical Physics, Niels Bohrweg 2, Leiden, 2333 CA, The Netherlands

(Received 16 July 2017; published 22 September 2017)

The loop extrusion theory assumes that protein factors, such as cohesin rings, act as molecular motors that extrude chromatin loops. However, recent single molecule experiments have shown that cohesin does not show motor activity. To predict the physical mechanism involved in loop extrusion, we here theoretically analyze the dynamics of cohesin rings on a loop, where a cohesin loader is in the middle and unloaders at the ends. Cohesin monomers bind to the loader rather frequently and cohesin dimers bind to this site only occasionally. Our theory predicts that a cohesin dimer extrudes loops by the osmotic pressure of cohesin monomers on the chromatin fiber between the two connected rings. With this mechanism, the frequency of the interactions between chromatin segments depends on the loading and unloading rates of dimers at the corresponding sites.

DOI: 10.1103/PhysRevE.96.030402

I. INTRODUCTION

Chromatin conformation capture and related techniques have shown that chromosomes in a cell nucleus form so-called topologically associated domains (TADs), contiguous regions of enriched contact frequency that are isolated from neighboring regions [1]. A recent theory predicts that TADs are composed of stochastic chromatin loops, which are produced by the loop extrusion mechanism; a protein factor, called a loop extrusion factor (LEF), acts as a molecular motor that extrudes a stretch of a chromatin fiber to form a loop and the loop size is limited by proteins, called boundary elements (BEs), that stop the motion of LEFs [2]. This theory predicts that the interactions of chromatin in TADs are local and stochastic, in agreement with experiments [1], for a window of parameters involved in the processivity of LEFs and the average distance between LEFs.

It has been proposed that cohesin, one of the SMC (structural maintenance of chromosomes) proteins, act as LEFs and CTCF acts as BEs [2]. While the roles of CTCF as BEs seem to be consistent with experiments [1,3], recent single molecule experiments showed that cohesin rings diffuse randomly along DNA, in contrast to the expectation of cohesin as a molecular motor [4–6]. Whether the diffusion constant of cohesin depends on the concentration of ATP is controversial [4–6]. These experiments also showed that a cohesin ring can contain only one chromatin fiber in its pore and thus must form dimers to mediate chromatin interactions. In view of these experimental results, how can cohesin rings extrude loops and ensure the locality and stochasticity of chromatin interactions, provided that cohesin rings indeed produce chromatin loops? Here we analyze the dynamics of cohesin rings that do not show motor activity on a chromatin fiber to predict a possible mechanism of the loop extrusion process.

II. MODEL

Single molecule experiments have shown that cohesin rings are (1) preferentially loaded to chromatin fibers at specific sites, occupied by loader proteins (such as Nipbl) and are (2) topologically bound to chromatin fibers for a relatively long time by using ATP [4,5]. In some cases, these proteins are (3) preferentially unloaded at another specific site, occupied by unloader proteins (such as Wapl) [7].

We treat a stretch of chromatin fiber as a 1D lattice of binding sites (of size a), which can be occupied by cohesin rings. The loading site is in the middle of the fiber stretch ($z = 0$) and the unloading sites are at its two ends ($z = \pm M$) (see Fig. 1). We assume that cohesin rings are usually loaded as monomers and are only occasionally loaded as dimers. Cohesin dimers transiently associate two segments of the fiber. Both cohesin monomers and dimers do not dissociate from the fiber until they reach the unloading site. The latter unloading condition is to simplify the model and is not essential. In the Supplemental Material (SM), we treat cases in which cohesin rings are unloaded on the track with a uniform rate [8]. The region in between the two unloading sites is bound by two BEs (such as two CTCF proteins of converging orientation) so that cohesin rings do not diffuse out from the region. Once loaded on the fiber, cohesin monomers cannot stick together to form dimers; this ensures the locality of the chromatin interactions (shown in Ref. [2]).

The flux J_m of cohesin monomers has the form

$$J_\mathrm{m} = -D_\mathrm{c}\frac{\partial}{\partial z}\psi_\mathrm{m}(z) - \frac{D_\mathrm{c}}{k_\mathrm{B}T}\psi_\mathrm{m}(z)\frac{\partial}{\partial z}\Pi_\mathrm{osm}(z), \qquad (1)$$

where $\psi_\mathrm{m}(z)$ is the probability of finding a cohesin monomer at the zth binding site ($-M < z < M$), D_c is the diffusion constant of cohesin rings, k_B is the Boltzmann constant, and T is the absolute temperature (see the SM for the derivation [8]). $\Pi_\mathrm{osm}(= -k_\mathrm{B}T\log[1-\psi_\mathrm{m}(z)])$ is the osmotic pressure generated by monomers [this returns to the ideal gas form for $\psi_\mathrm{m}(z) < 1$]. It has the dimension of energy because it acts parallel to the 1D lattice. The first term on the right-hand side of Eq. (1) is the thermal diffusion of

*tyamamoto@nuap.nagoya-u.ac.jp

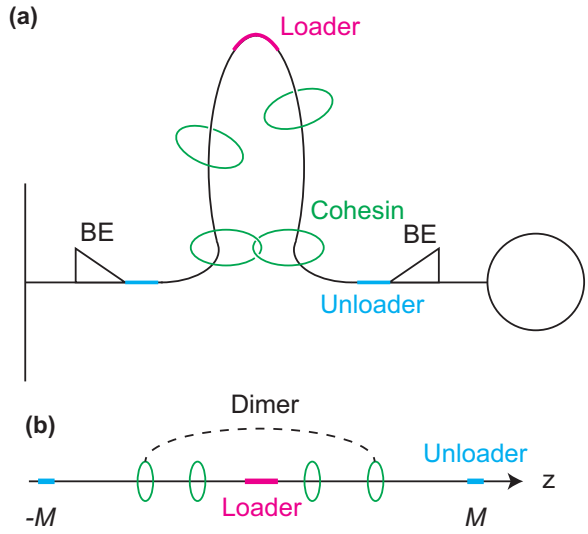

FIG. 1. (a) Chromatin fiber with a loading site for cohesin rings in the middle (magenta) and corresponding unloading sites at the ends (blue). Cohesin rings (green) topologically bind to the fiber as monomers or dimers. The two ends of this fiber stretch feature boundary elements (triangles) beyond which the cohesin rings cannot diffuse. (b) In our model we treat the chromatin fiber as a 1D lattice along which the cohesin monomers and dimers diffuse.

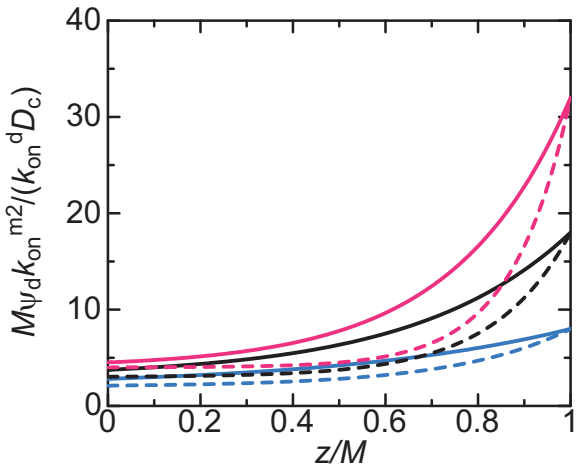

FIG. 2. The probability distribution $\psi_d(z)$ of dimers [rescaled by $k^d_{on}D_c/(k^{m2}_{on}M)$] is shown as a function of the rescaled position z/M for $Mk^m_{on}c/D_c = 3.4$ (cyan), 150.6 (black), and 1092.0 (magenta). We used $Mk^d_{off}/D_c = 0.5$ and $Mk^m_{off}/D_c = 5.0$. The solid and broken curves are derived by using Eq. (5) and Eq. (S12) in the Supplemental Material [8], respectively.

monomers and the second term is the flux generated by the osmotic pressure $\Pi_{osm}(z)$. We assumed that cohesin dimers bind to the chromatin fiber only rarely.

In steady state, J_m does not depend on the positions z and is equal to both the loading rate of cohesin monomers at the loading site, $k^m_{on}c[1-\psi_m(0)]$, and the unloading rate of these monomers at the unloading site, $k^m_{off}\psi_m(M)$ (k^m_{on} and k^m_{off} are the rate constants that account for the loading and unloading of monomers, respectively, and c is the concentrations of cohesin in solution). These boundary conditions highlight the fact that cohesin rings are topologically bound to the chromatin fiber for a relatively long time by using ATP; cohesin rings are unloaded at a site far away from the loading site, breaking the detailed balance. With these boundary conditions, Eq. (1) leads to the probability $\psi_m(z)$ in the form

$$\psi_m(z) = 1 - \left(1 - \frac{J_m}{k^m_{off}}\right)e^{J_m(z-M)/D_c}, \tag{2}$$

where the flux J_m is determined by the relationship

$$J_m = k^m_{on}c\left(1 - \frac{J_m}{k^m_{off}}\right)e^{-J_mM/D_c}. \tag{3}$$

The flux J_d of cohesin dimers has the form

$$J_d = -D_c\frac{\partial}{\partial z}\psi_d(z) - \frac{D_c}{k_BT}\psi_d(z)\frac{\partial}{\partial z}\Pi_{osm}(z), \tag{4}$$

where $\psi_d(z)$ is the probability of finding a cohesin dimer at z. Equation (4) is derived by assuming that the two rings of each dimer diffuse independently (see the SM for the derivation [8]) and by neglecting the contributions of the conformational entropy of the chromatin fiber (see also the SM [8]). The first term on the right-hand side of Eq. (4) is due to the thermal diffusion of cohesin dimers and the second term is due to the osmotic pressure $\Pi_{osm}(z)$ generated by cohesin monomers. For simplicity, we assumed that the diffusion constant of cohesin dimers is equal to the diffusion constant of cohesin monomers.

In steady state, the flux J_d does not depend on the position z and is equal to both the loading rate of dimers at the loading site, $k^d_{on}c^2[1-\psi_m(0)]^2$, and the unloading rate of dimers at the unloading site, $k^d_{off}\psi_d(M)$ (k^d_{on} and k^d_{off} are the rate constants that account for the loading and unloading of dimers, respectively). These boundary conditions highlight the fact that cohesin rings are topologically bound to the chromatin fiber for a relatively long time by using ATP; cohesin rings are unloaded at a site far away from the loading site, breaking the detailed balance. With these boundary conditions, Eq. (4) leads to the probability $\psi_d(z)$ in the form

$$\psi_d(z) = \frac{J_d}{J_m}(1 - \alpha e^{-J_m(M-z)/D_c}), \tag{5}$$

with the flux of dimers $J_d = k^d_{on}J^2_m/k^{m2}_{on}$ and introducing a factor $\alpha(= 1 - J_m/k^d_{off})$.

III. RESULTS AND DISCUSSION

The probability distribution $\psi_d(z)$ corresponds to the average frequency of contacts between the chromatin segments at z and $-z$. The distribution function $\psi_d(z)$ is a monotonically increasing function of the position z for $\alpha < 0$, in contrast to the prediction of simple thermal diffusion (see the solid curves in Fig. 2). This is because the motions of cohesin dimers are driven by the osmotic pressure generated by cohesin monomers [see the second term of Eq. (4)]. Hi-C experiments have shown that in many cases, the contact frequency of chromatin segments has a peak at the bases of the loop [1]. This implies that the probability $\psi_d(z)$ is a steeply increasing function of the position z; the factor α is negative and the length scale $aD_c/(J_mM)$ of dimer distribution (a is the size of a binding site, which we use for consistency, but does

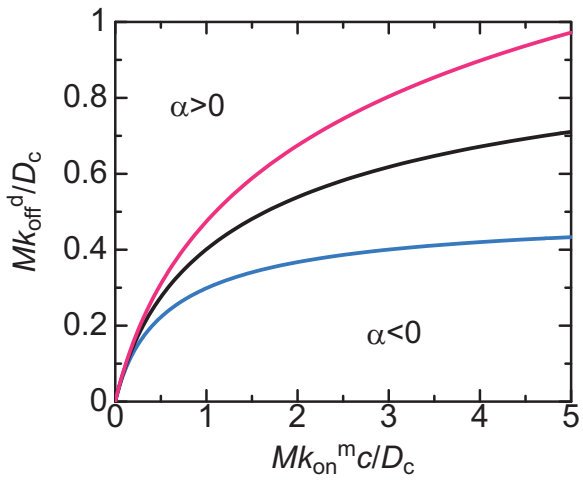

FIG. 3. Diagram of states defined by the sign of the factor α [see Eq. (5)] is shown as a function of the rescaled monomer loading rate $k_{on}^{m}cM/D_c$ and the rescaled dimer unloading rate $k_{off}^{d}M/D_c$. The boundary is shown for the values of the rescaled monomer unloading rate $k_{off}^{m}M/D_c = 0.5$ (blue), 1.0 (black), and 2.0 (magenta).

not matter in the following calculations) is smaller than the order of 3.4 μm ($\simeq$10 kbps). This corresponds to cases in which the loading and unloading rates k_{on}^{d} and k_{off}^{d} of dimers is relatively small and the loading and unloading rates k_{on}^{m} and k_{off}^{m} of monomers is large [see Eq. (5)]. More quantitatively, for $a^2 D_c \sim 1$ μm^2/min [4,6] and $aM \sim 340$ μm ($\simeq 1$ Mbps) [9], the monomer flux J_m must be larger than the order of 10^{-3} min^{-1} so that $aD_c/(J_m M) < 3.4$ μm ($\simeq$10 kbps) and the unloading rate k_{off}^{d} of dimers must be smaller than the monomer flux. Our theory predicts that cohesin rings extrude loops, even without motor activity, due to the osmotic pressure generated by monomers. The stochasticity of chromatin interactions in TADs is ensured by the fact that dimers are loaded to the chromatin fiber only occasionally.

Our theory predicts that the distribution function $\psi_d(z)$ of cohesin dimers changes from a monotonically decreasing function, $\alpha > 0$, to a monotonically increasing function, $\alpha < 0$, with increasing the concentration c of cohesin in the solution (see Fig. 3). This is because the flux of monomers increases with increasing the concentration c [see Eq. (3) and below Eq. (5)]. The peaks of the probability distributions of monomers and dimers, characterized by the length scale $aD_c/(J_m M)$, become steeper as the concentration c increases. The contact probability of chromatin segments, which scales as J_d/J_m, increases with increasing the concentration c. These predictions may be accessible by single molecule experiments.

Recent experiments have shown that the diffusion constant of cohesin rings changes with post-translational modifications and/or accessory proteins [6]. Our theory predicts that the parameter α decreases with increasing the diffusion constant D_c and that the width $aD_c/(J_m M)$ of the peak of the distribution function $\psi_d(z)$ at the loop base ($z = M$) increases with the diffusion constant D_c. These predictions may be experimentally accessible by using post-translational modifications and/or accessory proteins (for cases in which the modification changes the diffusion constants of monomers and dimers in a similar manner).

Our theory predicts that cohesin dimers extrude chromatin loops due to the osmotic pressure of cohesin monomers, if (1) both cohesin monomers and dimers are preferentially loaded at the loading site and (2) they are loaded on chromatin for a relatively long time by using ATP. The chromatin interactions are stochastic and local, as predicted by the loop extrusion theory, if (3) cohesin dimers are loaded to the chromatin fiber only rarely and (4) cohesin monomers do not assemble into dimers after they are loaded onto the chromatin fiber. This is a possible mechanism behind the loop extrusion process, provided that cohesin rings indeed extrude chromatin loops. However, it cannot be excluded at this time that other molecules actively extrude loops or that cohesin shows motor activity in an unidentified condition. Indeed, recent experiments indicate that condensin, another SMC protein, has motor activity [10]. If their motion is driven by the motor activity, the average speed of cohesin rings does not depend on the position z along the chromatin fiber (provided that the fiber is ideally uniform). Experimentally measuring the speed of cohesin rings as a function of the position z may provide useful information about the physical mechanisms involved in the loop extrusion process.

We used a couple of assumptions to simplify the model. First, cohesin rings are loaded on the chromatin fiber until they are released by unloader proteins. Second, the conformational entropy of the chromatin fiber does not drive the flux of cohesin dimers. In the SM, we show that releasing these assumptions makes the gradient of the distribution function $\psi_d(z)$ (with respect to the position z) zero at $z = M$ (because BEs stop the motion of cohesin monomers and dimers) and produces a peak at the loading site, but does not change the physics (see Secs. S2 and S3 in the SM [8]). Third, the two rings of each dimer diffuse independently. If the two rings diffuse in registry, the dimer diffusion constant is one-half times the monomer diffusion constant [see Eq. (S12) in the Supplemental Material [8]]. Then, the peak of the distribution function $\psi_d(z)$ at $z = M$ is sharper than in the case where the two rings of each dimer diffuse independently (see the broken curves in Fig. 2). Fourth, we assumed that the diffusion constants of monomers and dimers are equal. However, cohesin monomers can simply diffuse along the chromatin fiber, whereas the conformational dynamics of the chromatin fiber may be involved in the diffusion of cohesin dimers. If this is the case, the diffusion constant of dimers is equal to the Rouse diffusion constant [11] and is much smaller than the diffusion constant of monomers. The peak width of the contact frequency at the loop base is thus much steeper than our prediction. Finally, our theory is a mean-field theory and thus may not capture the correlation involved in the fact that monomers cannot penetrate through dimer rings and vice versa. We expect that the latter correlation is not significant for cases in which the diffusion constants of monomers and dimers are approximately equal.

Note added. Recently, we became aware of an arXiv preprint, where it is also proposed that the loop extrusion might be driven by an osmotic mechanism [12]. Unlike in our model, there are only dimers which produce loops and at the same time generate the osmotic pressure. If many cohesin dimers are loaded on a chromatin loop so that the osmotic pressure becomes significant, the loop may be tightly bound by the dimers, in contrast to the fact that the contact

frequencies in TADs are only moderate, with the exception of the peak loci at the TAD boundaries [1,2]. If there are only a few dimers, the osmotic pressure may not be significant enough to produce a peak of the contact frequency at the TAD boundaries. The window of parameters with which the contact frequency shows a peak at the loop base may thus be rather limited or even may not exist if the osmotic pressure of dimers is the only mechanism to produce the peak. In contrast, by dividing the tasks of osmotic pressure generation and loop formation to monomers and dimers, we predict that the interaction frequency shows a peak at the loop base in agreement with experiments [1], while the stochasticity and locality of chromatin interactions are ensured. This theory singles out the contributions of osmotic pressure and shows quantitative predictions. These predictions may be useful to elucidate the physical mechanism that stabilizes TADs.

[1] S. S. P. Rao, M. H. Huntley, N. C. Durand, E. K. Stamenova, I. D. Bochkov, J. T. Robinson, A. L. Sanborn, I. Machol, A. D. Omer, E. S. Lander, and E. Lieberman Aiden, Cell **159**, 1665 (2014).

[2] G. Fudenberg, M. Imakaev, C. Lu, A. Goloborodko, N. Abdennur, and L. A. Mirny, Cell Rep. **15**, 2038 (2016).

[3] C. Hou, H. Zhao, K. Tanimoto, and A. Dean, Proc. Natl. Acad. Sci. USA **105**, 20398 (2008).

[4] J. Stigler, G. Ö. Camdere, D. E. Koshland, and E. C. Greene, Cell Rep. **15**, 988 (2016).

[5] I. F. Davidson, D. Goetz, M. P. Zaczek, M. I. Molodtsov, P. J. Huis in't Veld, F. Weissman, G. Litos, D. A. Cisneros, M. Ocampo-Hafalla, R. Ladurner, F. Uhlmann, A. Vaziri, and J.-M. Peters, EMBO J. **35**, 2671 (2016).

[6] M. Kanke, E. Tahara, P. J. Huis in't Veld, and T. Nishiyama, EMBO J. **35**, 2686 (2016).

[7] G. A. Busslinger, R. R. Stocsits, P. van der Lelij, E. Axesson, A. Tedeschi, N. Galjart, and J. M. Peters, Nature (London) **544**, 503 (2017).

[8] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevE.96.030402 for derivations and additional material.

[9] J. E. Phillips-Cremins, M. E. G. Sauria, A. Sanyal, T. I. Gerasimova, B. R. Lajoie, J. S. K. Bell, C. T. Ong, T. A. Hookway, C. Guo, Y. Sun, M. J. Bland, W. Wagstaff, S. Dalton, T. C. McDevitt, R. Sen, J. Dekker, J. Taylor, and V. G. Corces, Cell **153**, 1281 (2013).

[10] T. Terekawa, S. Bisht, J. M. Eeftens, C. Dekker, C. H. Haering, and E. C. Greene, Science (2017), doi: 10.1126/science.aan6516.

[11] M. Doi and S. F. Edwards, *The Theory of Polymer Dynamics* (Oxford University Press, Oxford, 1986).

[12] C. A. Brackley, J. Johnson, D. Michieletto, A. N. Morozov, M. Nicodemi, P. R. Cook, and D. Marenduzzo, arXiv:1612.07256v1.