Extrusion of chromatin loops by a composite loop extrusion factor

Hao Yan,[1,2] Ivan Surovtsev,[2,3] Jessica F Williams,[3] Mary Lou P Bailey,[1,4] Megan C King,[1,3,5] and Simon G J Mochrie[1,2,4,*]

[1] *Integrated Graduate Program in Physical and Engineering Biology, Yale University, New Haven, Connecticut 06511, USA*
[2] *Department of Physics, Yale University, New Haven, Connecticut 06511, USA*
[3] *Department of Cell Biology, Yale School of Medicine, New Haven, Connecticut 06520, USA*
[4] *Department of Applied Physics, Yale University, New Haven, Connecticut 06511, USA*
[5] *Department of Molecular, Cellular, and Developmental Biology, Yale University, New Haven, Connecticut 06511, USA*

(Dated: July 12, 2021)

Chromatin loop extrusion by Structural Maintenance of Chromosome (SMC) complexes is thought to underlie intermediate-scale chromatin organization inside cells. Motivated by a number of experiments suggesting that nucleosomes may block loop extrusion by SMCs, such as cohesin and condensin complexes, we introduce and characterize theoretically a composite loop extrusion factor (composite LEF) model. In addition to an SMC complex that creates a chromatin loop by encircling two threads of DNA, this model includes a remodeling complex that relocates or removes nucleosomes as it progresses along the chromatin, and nucleosomes that block SMC translocation along the DNA. Loop extrusion is enabled by SMC motion along nucleosome-free DNA, created in the wake of the remodeling complex, while nucleosome re-binding behind the SMC acts as a ratchet, holding the SMC close to the remodeling complex. We show that, for a wide range of parameter values, this collection of factors constitutes a composite LEF that extrudes loops with a velocity, comparable to the velocity of remodeling complex translocation on chromatin in the absence of SMC, and much faster than loop extrusion by an isolated SMC that is blocked by nucleosomes.

I. INTRODUCTION

Exquisite spatial organization is a defining property of chromatin, allowing the genome both to be accommodated within the volume of the cell nucleus, and simultaneously accessible to the transcriptional machinery, necessary for gene expression. On the molecular scale, histone proteins organize 147 bp of DNA into nucleosomes, that are separated one from the next by an additional 5-60 bp [1]. On mesoscopic scales ($10^5 - 10^6$ bp), it has long been understood that loops are an essential feature of chromatin organization. The recent development of chromosome conformation capture (Hi-C) techniques now enables quantification of chromatin organization via a proximity ligation assay, that yields a map of the relative probability that any two genomic locations are in contact with each other [2]. Hi-C contact maps have led to the identification of topologically associating domains (TADs) as fundamental elements of intermediate-scale chromatin organization [3–7]. Genomic regions inside a TAD interact frequently with each other, but have relatively little contact with regions in even neighboring TADs.

Although how TADs arise remains uncertain, the loop extrusion factor (LEF) model has emerged as the preferred candidate mechanism for TAD formation. In this model, LEFs – identified as the Structural Maintenance of Chromosome (SMC) complexes, cohesin and condensin – encircle two chromatin threads, forming the base of a loop, and then initiate loop extrusion [8–13]. Efficient topological cohesin loading onto chromatin, as envisioned by the LEF model, depends both on the presence of the Scc2-Scc4 cohesin loading complex and on cohesin's ATPase activity [14]. Loop extrusion proceeds until the LEF is blocked by another LEF or until it encounters a boundary element, generally identified as DNA-bound CCCTC-binding factor (CTCF), or until it dissociates, causing the corresponding loop to dissipate. Thus, a population of LEFs leads to a dynamic steady-state chromatin organization. As may be expected, based on the correlation between TAD boundaries and CTCF binding sites [15], this model recapitulates important features of experimental Hi-C contact maps [8, 10, 11].

The LEF model was recently bolstered by beautiful single-molecule experiments that directly visualized DNA loop extrusion by condensin [16] and cohesin [17]. However, both of these studies focused on the behavior of the SMC complex on naked DNA, whereas inside cells DNA is densely decorated with nucleosomes. Ref. [17] (and then Ref. [18]) did also show that cohesin could compact lambda DNA (48,000 bp) loaded with about three nucleosomes, but this nucleosome density (6×10^{-5} bp^{-1}) is nearly 100-fold less than the nucleosome density in chromatin (5×10^{-3} bp^{-1}).

The notion that nucleosomes might actually represent a barrier for SMC translocation and therefore loop extrusion is suggested by measurements that reveal that cohesin motions on nucleosomal DNA are much reduced compared to those on naked DNA [19, 20]. Further supporting the hypothesis that nucleosomes hinder SMC-driven loop extrusion are several studies indicating that

* simon.mochrie@yale.edu

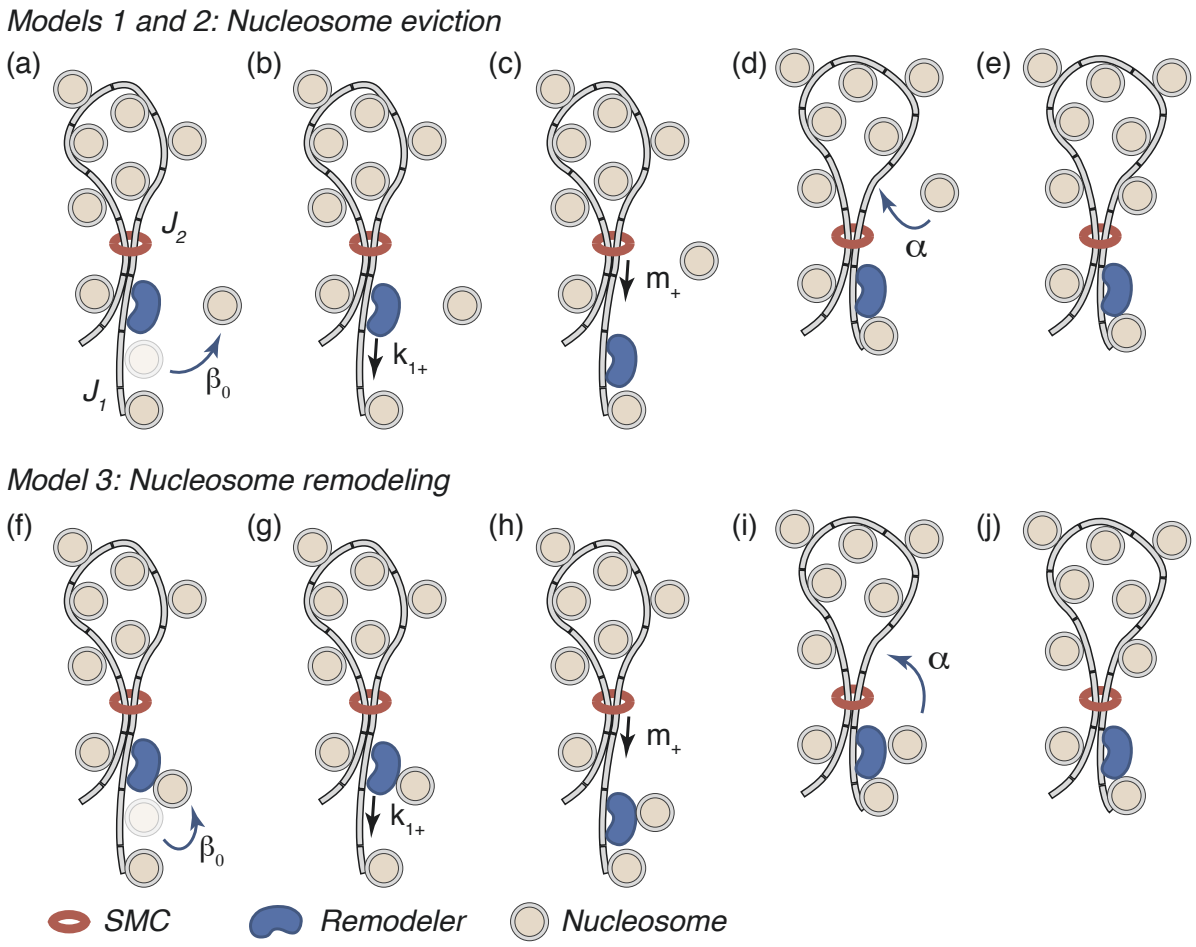

FIG. 1. Loop extrusion via a composite LEF, comprising an SMC complex, which forms a ring around two nucleosome-free sections of DNA, nucleosomes that block SMC translocation, and a remodeling complex which removes nucleosomes in front of the SMC. In our model, a single loop extrusion step starts when the remodeling complex forces a nucleosome from the DNA ahead of the remodeler, thus moving the junction (J1) between nucleosomal DNA and naked DNA one step forward. β_0 is the rate of nucleosome dissociation (a) or remodeling (f) when the remodeler is next to a nucleosome. Next, the remodeler moves into the resultant nucleosome-free region, (b) and (g). k_{1+} is the rate at which the remodeler steps forward, when the remodeler-nucleosome separation is one step. Then, the SMC complex moves into the new nucleosome-free region left behind the remodeler (c) and (h). m_+ is the rate at which the the SMC steps forward on nucleosome-free DNA. Finally, a nucleosome rebinds behind the SMC complex, moving the second junction (J2) between nucleosomal DNA and naked DNA one step forward, and so preventing the SMC from subsequently backtracking. α is the rate of nucleosome rebinding (d) or re-formation (i). After these four sub-steps, the LEF configuration is the same as before the first step, but the loop is one step larger, (e) and (j). The top row (a-e) illustrates a hypothetical scenario (models 1 and 2) in which the displaced nucleosome is in solution before rebinding DNA behind the SMC. The bottom row (f-j) illustrates an alternative "remodeled-nucleosome" scenario (model 3) in which the displaced nucleosome remains associated with the remodeling complex before rebinding DNA behind the SMC.

cohesin translocation requires transcription-coupled nucleosome remodeling [20–24]. In particular, Ref. [20] demonstrates that cohesin, recruited to one genomic location by a cohesin loading complex, is relocated to another by RNA polymerase (Pol II) during transcription. Finally, Ref. [25] found that presence of nucleosomes in *Xenopus laevis* egg extract prevented DNA exposed to the extract from looping and compaction.

In this paper, motivated by the possibility that nucleosomes block loop extrusion by SMCs, we introduce and characterize theoretically a composite loop extrusion factor (composite LEF) model that realizes chromatin loop extrusion. Our model focuses specifically on composite LEF translocation and the process of ongoing extrusion of chromatin loops. Nevertheless, we envision that loop extrusion terminates when a composite LEF encounters a boundary element, such as CTCF, just as in existing LEF models. In addition, to ensure that composite LEFs grow loops (rather than shrink them), we also suppose that remodelers are recruited to SMCs with a definite orien-

tation. However, consideration of the molecular mechanisms by which composite LEF loop extrusion might be terminated by CTCF, or initiated, lies beyond the scope of this paper.

Fig. 1 is a cartoon representation of the composite LEF model. As illustrated in the figure, in addition to an SMC complex that encircles two threads of DNA, creating a chromatin loop, the model includes a remodeling complex, that removes or relocates nucleosomes as it translocates along chromatin, and nucleosomes, that create a barrier for SMC motion. Both the remodeler and the nucleosomes are essential components of the composite LEF. We envision that when the ring-like SMC complex is threaded by DNA, it can move along the DNA until it encounters a nucleosome, which blocks its motion. We hypothesize that the SMC's ATPase activity does not exert enough force to move a nucleosome, even though it may give rise to directional loop extrusion on naked DNA. Without nucleosome remodeling, therefore, an SMC complex remains trapped by its surrounding nucleosomes at a more-or-less fixed genomic location. Loop extrusion is enabled by SMC motion along the nucleosome-free thread of DNA, that is created in the wake of the remodeling complex, and is maintained by the SMC being held close to the remodeling complex by the ratcheting action of nucleosomes re-locating to behind the SMC. The composite LEF, illustrated in Fig. 1, extrudes the right-hand thread of the chromatin loop, embraced by its constituent SMC complex. The left-hand thread of the loop remains encircled by the SMC at a fixed genomic location, with the SMC trapped by its surrounding nucleosomes. In our model, two-sided loop extrusion would require a remodeler on each thread, translocating in opposite directions. The model is agnostic concerning the specific identity of the remodeler, except that it must be able to displace nucleosomes or alter their configuration in a manner that allows the SMC to subsequently pass them by. The top row of Fig. 1 illustrates a hypothetical process, in which the displaced nucleosome unbinds from ahead of the remodeler, before the same or a different nucleosome subsequently rebinds behind the SMC. The bottom row illustrates an alternative version of the model, in which the displaced nucleosome remains associated with the LEF in a transient, "remodeled" configuration, that is permissive to loop extrusion.

This paper is organized as follows. In Sec. II, we calculate the velocities of one-sided loop extrusion for three, slightly different versions of the composite LEF model. In fact, differences among the loop extrusion velocities of the different models are small. In Sec. III, we examine the results of Sec. II to elucidate the conditions required for efficient loop extrusion. We also compare the composite LEF's loop extrusion velocity to the velocities of the remodeler and the SMC, each translocating alone on chromatin. For a broad range of parameter values, we find that the model's component factors can indeed be sensibly identified as a composite LEF, that can extrude chromatin loops at a velocity that is comparable to that of isolated remodeler translocation on chromatin, and much faster than loop extrusion by an isolated SMC, that is blocked by nucleosomes. Finally, in Sec. IV, we conclude.

II. THEORY

The results presented in this section rely on and were guided by the calculations and ideas presented in Refs. [26–28], concerning other examples of biological Brownian ratchets. To calculate the loop extrusion velocity, v, in terms of the rates of remodeling complex forward (k_+) and backward (k_-) stepping on DNA, the rates of SMC forward (m_+) and backward (m_-) stepping on DNA, and the rates of nucleosome binding (α) and unbinding (β), *etc.*, we make a number of simplifying assumptions. First, we consider chromatin as a sequence of nucleosome binding sites. Second, we assume that the none of the SMC complex, the remodeling complex, and nucleosomes can occupy the same location, *i.e.* we assume an infinite hard-core repulsion between these factors, that prevents their overlap. Third, we assume that there are well-defined junctions between bare DNA and nucleosomal DNA in front of the remodeler (junction 1) and behind the SMC loop (junction 2), so that when a remodeler forces a nucleosome from junction 1, subsequently it relocates to junction 2. Finally, we hypothesize that, although SMCs can not push nucleosomes out of their way, the remodeling complex can. Following Refs. [27, 28], we actualize this nucleosome-ejecting activity via a nearest-neighbor repulsive interaction, ΔG, between the remodeling complex and junction 1. For convenience, the key parameters of our models are summarized in Table I.

A. Model 1

First, we consider a streamlined model (model 1), which assumes that the nucleosome unbinding and rebinding rates are much faster than the remodeling complex and SMC forward- and backward-stepping rates. Because of this separation of time scales, we can consider that the SMC and remodeler move in a free energy landscape defined by the time-averaged configuration of nucleosomes [26]. Thus, when the remodeling complex and junction 1 are next to each other (zero separation), the free energy is ΔG, corresponding to the nearest-neighbor remodeler-junction repulsive interaction, or, when there are n nucleosome binding sites between the remodeling complex and junction 1, the free energy is $n\Delta g$, corresponding to the free energy of n unbound nucleosomes in front of the remodeling complex. A straightforward equilibrium statistical mechanical calculation then informs us that the probability that the remodeling complex and

TABLE I. List of symbols representing the key parameters used in this paper. All quantities are referred to a lattice of nucleosome binding sites

Symbol	Quantity
Δg	Nucleosome binding free energy[a]or remodeled nucleosome free energy.[b]
ΔG	Free energy of nucleosome-remodeler nearest-neighbor repulsion.
ΔG_R	Free energy change when the remodeler steps forward.
ΔG_S	Free energy change when the SMC complex steps forward.
P_1	Probability that the remodeler and junction 1 are not next to each other.
P_2	Probability that the SMC complex and junction 2 are not next to each other.
P_3	Probability that the remodeler and the SMC complex are not next to each other.
k_+	Remodeler forward stepping rate on DNA away from nucleosomes.
k_-	Remodeler backward stepping rate on DNA away from nucleosomes.
k_{1+}	Remodeler forward stepping rate on DNA into nucleosome contact.
k_{0-}	Remodeler backward stepping rate on DNA out of nucleosome contact.
m_+	SMC complex forward stepping rate on DNA.
m_-	SMC complex backward stepping rate on DNA.
α	Nucleosome binding rate away from the remodeler.
β	Nucleosome unbinding rate away from the remodeler.
α_1	Nucleosome binding rate into remodeler contact.
β_0	Nucleosome unbinding rate from remodeler contact.
b	Step size, taken to be the nucleosome separation.

[a] Models 1 and 2.
[b] Model 3

junction 1 are not next to each other is

$$P_1 = \frac{1}{1 + e^{\frac{\Delta g - \Delta G}{k_B T}} - e^{-\frac{\Delta G}{k_B T}}}. \quad (1)$$

Similarly, the probability that the SMC and junction 2 are not next to each other is

$$P_2 = e^{-\frac{\Delta g}{k_B T}}, \quad (2)$$

because we assume there is not a SMC-nucleosome nearest-neighbor interaction, beyond the requirement that they not be at the same location.

The principle of detailed balance informs us that the ratio of forward and backward transition rates are given by a Boltzmann factor. Therefore, when the the remodeling complex and junction 1 are not next to each other, we expect

$$\frac{k_+}{k_-} = e^{-\frac{\Delta G_R}{k_B T}}, \quad (3)$$

where ΔG_R is the free energy change involved in moving the remodeler one step forward. However, when the remodeling complex and junction 1 are next to each other, this ratio of rates is modified, because of the nucleosome-remodeling complex repulsion:

$$\frac{k_{1+}}{k_{0-}} = e^{-\frac{\Delta G}{k_B T}} \frac{k_+}{k_-}, \quad (4)$$

where k_{1+} is the remodeling complex forward stepping rate, when the remodeler-junction 1 separation is one step, and k_{0-} is the remodeling complex backward stepping rate when the remodeler-junction 1 separation is zero. As discussed in detail in Refs. [27, 28], to satisfy Equation (4), in general, we can write

$$k_{1+} = e^{-\frac{\Delta G f}{k_B T}} k_+, \quad (5)$$

$$k_{0-} = e^{\frac{\Delta G (1-f)}{k_B T}} k_-, \quad (6)$$

where $0 < f < 1$ [27, 28]. However, as discussed in detail in Refs. [27] and [28] in an analogous context, the choice $f = 0$ maximizes the composite LEF velocity. Therefore, we pick $f = 0$, so that

$$k_{1+} = k_+ \quad (7)$$

and

$$k_{0-} = e^{\frac{\Delta G}{k_B T}} k_-, \quad (8)$$

which satisfy Equation (4). Then, the mean velocity of the remodeling complex may be written

$$v_R = b k_+ P_1 - b k_- P_3 P_1 - b k_- e^{\frac{\Delta G}{k_B T}} P_3 (1 - P_1) \quad (9)$$

where P_3 is the probability that the remodeling complex and the SMC are not next to each other, and b is the step size along the DNA, taken to be the separation between nucleosomes for simplicity. The first term on the right-hand side of Equation (9) corresponds to stepping forward, which can only happen if the remodeling complex and junction 1 are not next to each other. The second term on the right-hand side of Equation (9) corresponds to stepping backwards in the case that the remodeling complex and junction 1 are not next to each other and the remodeling complex and the SMC complex are not

next to each other, in which case the rate of this process is k_-. The third term on the right-hand side of Equation (9) corresponds to stepping backwards in the case that the remodeling complex and junction 1 are next to each other and the remodeling complex and the SMC complex are not next to each other, in which case the rate of this process is $k_- e^{\frac{\Delta G}{k_B T}}$, according to Equation (8). Using Equation (1) in Equation (9), we find

$$v_R = b\frac{k_+ - k_- e^{\frac{\Delta g}{k_B T}} P_3}{1+e^{\frac{\Delta g - \Delta G}{k_B T}} - e^{-\frac{\Delta G}{k_B T}}}. \quad (10)$$

We can also calculate the diffusivity of the remodeler:

$$D_R = \tfrac{1}{2}\left(b^2 k_+ P_1 + b^2 k_- P_3 P_1 + b^2 k_- e^{\frac{\Delta G}{k_B T}} P_3(1-P_1)\right)$$
$$= \tfrac{1}{2} b^2 \frac{k_+ + k_- e^{\frac{\Delta g}{k_B T}} P_3}{1+e^{\frac{\Delta g - \Delta G}{k_B T}} - e^{-\frac{\Delta G}{k_B T}}}. \quad (11)$$

Similar reasoning informs us that the velocity and diffusivity of the SMC complex are

$$v_S = b(m_+ P_3 - m_- P_2) = b(m_+ P_3 - m_- e^{-\frac{\Delta g}{k_B T}}), \quad (12)$$

and

$$D_S = \frac{1}{2} b^2 (m_+ P_3 + m_- e^{-\frac{\Delta g}{k_B T}}), \quad (13)$$

respectively.

Equation (10) shows that the velocity of the remodeling complex, v_R, decreases with increasing P_3, while Equation (12) shows that the velocity of the SMC complex, v_S, increases with increasing P_3. To realize a composite LEF, P_3 must take on a value that causes these two velocities to coincide, so that the remodeling complex and the SMC complex translocate together with a common velocity, v, given by $v = v_R = v_S$. Equations (10) and (12) constitute two equations for the two unknowns, P_3 and v. Solving yields

$$P_3 = \frac{\frac{k_+}{k_- m_+} + (e^{-\frac{\Delta g}{k_B T}} - e^{-\frac{\Delta g + \Delta G}{k_B T}} + e^{-\frac{\Delta G}{k_B T}})\frac{m_-}{k_- m_+}}{\frac{e^{\frac{\Delta g}{k_B T}}}{m_+} + \frac{1+e^{\frac{\Delta g - \Delta G}{k_B T}} - e^{-\frac{\Delta G}{k_B T}}}{k_-}} \quad (14)$$

and

$$v = b\frac{\frac{k_+}{k_-} - \frac{m_-}{m_+}}{\frac{e^{\frac{\Delta g}{k_B T}}}{m_+} + \frac{1+e^{\frac{\Delta g - \Delta G}{k_B T}} - e^{-\frac{\Delta G}{k_B T}}}{k_-}}. \quad (15)$$

Using this value for P_3, it further follows that

$$D_R = b^2 \frac{k_+}{1+e^{\frac{\Delta g - \Delta G}{k_B T}} - e^{-\frac{\Delta G}{k_B T}}} + \tfrac{1}{2} b^2 \frac{\frac{m_-}{m_+} - \frac{k_+}{k_-}}{\frac{e^{\frac{\Delta g}{k_B T}}}{m_+} + \frac{1+e^{\frac{\Delta g - \Delta G}{k_B T}} - e^{-\frac{\Delta G}{k_B T}}}{k_-}} \quad (16)$$

and

$$D_S = \tfrac{1}{2} b^2 \frac{\frac{k_+}{k_-} + (e^{-\frac{\Delta g}{k_B T}} - e^{-\frac{\Delta g + \Delta G}{k_B T}} + e^{-\frac{\Delta G}{k_B T}})\frac{m_-}{k_-}}{\frac{e^{\frac{\Delta g}{k_B T}}}{m_+} + \frac{1+e^{\frac{\Delta g - \Delta G}{k_B T}} - e^{-\frac{\Delta G}{k_B T}}}{k_-}} + \tfrac{1}{2} b^2 e^{-\frac{\Delta g}{k_B T}} m_-. \quad (17)$$

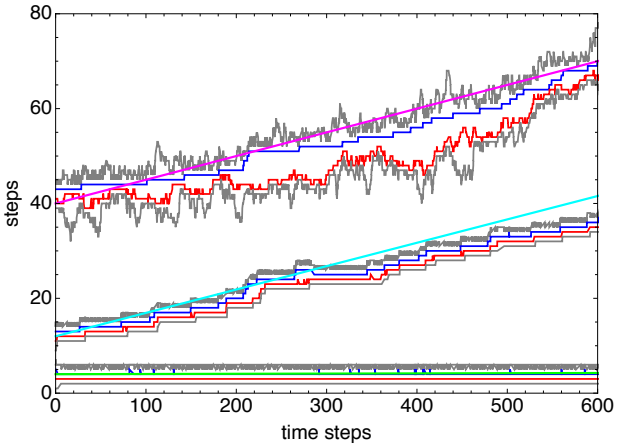

FIG. 2. Three example composite LEF trajectories from model 2 simulations. In each case, the positions versus time of the nucleosome junctions are shown gray, the remodeling complex is shown blue, and the SMC complex is shown red. When tracking together, each such group of four traces constitutes a composite LEF. The model parameters are $k_+ = 0.05$ per time step, $k_- = 5 \times 10^{-7}$ per time step, $m_+ = m_- = 0.3$ per time step, $\Delta G = 18.0 k_B T$, $\alpha = 1$ per time step, and $\beta - \alpha e^{\frac{-\Delta g}{k_B T}}$ for all three composite LEFs, but $\Delta g = 18.0 k_B T$ for the bottom group of traces, $\Delta g = 9.0 k_B T$ for the middle group of traces, and $\Delta g = 0.5 k_B T$ for the top group of traces. The cyan, green, and magenta lines each have a slope given by the theoretical composite LEF velocity for the parameters of each simulation.

B. Model 2

At the cost of a little complication, it is possible to calculate the composite LEF velocity, even when the nucleosome binding (α) and unbinding (β) rates are not much larger than k_+, k_-, m_+, and m_-. This model (model 2) is preferable *a priori* because we expect the nucleosome unbinding rate, β, to be small. In fact, the results obtained with model 2 are very similar to those obtained with model 1.

Similar to the remodeling complex forward- and backward-stepping rates, when the remodeling complex and junction 1 are adjacent, the nucleosome binding and unbinding rates are modified as follows:

$$\frac{\alpha_1}{\beta_0} = e^{-\frac{\Delta G}{k_B T}} \frac{\alpha}{\beta}, \quad (18)$$

where α_1 is the nucleosome binding rate when the remodeler-junction 1 separation is one step, and β_0 is the nucleosome unbinding rate when the remodeling complex and junction 1 are adjacent (separation 0). To satisfy Equation (18), we can write

$$\alpha_1 = e^{-\frac{\Delta G f}{k_B T}} \alpha, \quad (19)$$

and

$$\beta_0 = e^{\frac{\Delta G (1-f)}{k_B T}} \beta, \quad (20)$$

which stand alongside Equations (7) and (8). As above, we again choose $f = 0$, so that

$$\alpha_1 = \alpha \tag{21}$$

and

$$\beta_0 = e^{\frac{\Delta G}{k_B T}}\beta. \tag{22}$$

To proceed in this case, we first write down the mean velocity of junction 1:

$$v_{J1} = b(\beta - \alpha)P_1 + b\beta e^{\frac{\Delta G}{k_B T}}(1 - P_1), \tag{23}$$

where P_1 is the probability that the remodeling complex and junction 1 are not next to each other and b is the step size. Similarly, we can also write down the mean velocity of the remodeling complex:

$$v_R = b(k_+ - k_- P_3)P_1 - bk_- e^{\frac{\Delta G}{k_B T}}(1 - P_1)P_3, \tag{24}$$

where P_3 is the probability that the remodeling complex and the SMC complex are not adjacent to each other. Next, we write down the velocity of the SMC complex:

$$v_S = b(m_+ P_3 - m_- P_2), \tag{25}$$

where P_2 is the probability that the SMC complex and the junction between bare DNA and nucleosomal DNA behind the SMC complex, namely junction 2, are not adjacent to each other. Finally, we can write down the mean velocity of junction 2:

$$v_{J2} = b(\alpha P_2 - \beta). \tag{26}$$

For the composite LEF to translocate as a single entity, it is necessary for each of its component parts to translocate with a common velocity, v, where

$$v = v_{J1} = v_R = v_S = v_{J2}. \tag{27}$$

Solving Equations (23) through (27) for the four unknowns, namely v, and the probabilities, P_1, P_3, and P_2, yields the values of these quantities. To this end, first we solve Equations (23) and (24), assuming that junction 1 and the remodeling complex have a common velocity (v_1) with the result that

$$v_1 = b\frac{(\frac{k_+}{k_-} - \frac{\alpha}{\beta}P_3)}{(1 - e^{-\frac{\Delta G}{k_B T}})(\frac{1}{k_-} + \frac{P_3}{\beta}) + \frac{\alpha + k_+}{\beta k_-}e^{-\frac{\Delta G}{k_B T}}}. \tag{28}$$

Next, we solve Equations (25) and (26), assuming that the SMC complex and junction 2 have a common velocity (v_2). In this case, we find

$$v_2 = b\frac{\frac{m_+ P_3}{m_-} - \frac{\beta}{\alpha}}{\frac{1}{m_-} + \frac{1}{\alpha}} \tag{29}$$

for the SMC-junction 2 velocity. For these two pairs to translocate together, manifesting a four component, composite LEF, it is necessary that they share a common velocity, v, given by $v = v_1 = v_2$. Setting Equation (28) equal to Equation (29) and solving for P_3, we find

$$\begin{aligned} P_3 =& \frac{1}{2m_+(1 - e^{-\frac{\Delta G}{k_B T}})}\left(-\alpha - m_- - \frac{m_+}{k_-}(\alpha + k_+)e^{-\frac{\Delta G}{k_B T}} + (\frac{\beta m_-}{\alpha} - \frac{\beta m_+}{k_+})(1 - e^{-\frac{\Delta G}{k_B T}})\right) \\ &+ \frac{\beta m_-}{2m_+(1 - e^{-\frac{\Delta G}{k_B T}})}\sqrt{\begin{aligned}&\frac{4m_+}{\beta m_-}(1 - e^{-\frac{\Delta G}{k_B T}})(\frac{k_+}{\alpha k_-} + \frac{k_+}{k_- m_-} + \frac{\beta}{\alpha k_-}(1 - e^{-\frac{\Delta G}{k_B T}}) + (\frac{k_+}{\alpha k_-} + \frac{1}{k_-})e^{-\frac{\Delta G}{k_B T}}) \\ &+ \left(\frac{1}{\beta} + \frac{\alpha}{\beta m_-} + (\frac{\alpha m_+}{\beta k_- m_-} + \frac{k_+ m_+}{\beta k_- m_-})e^{-\frac{\Delta G}{k_B T}} + (\frac{m_+}{k_- m_-} - \frac{1}{\alpha})(1 - e^{-\frac{\Delta G}{k_B T}})\right)^2.\end{aligned}} \end{aligned} \tag{30}$$

The velocity of the composite LEF can be calculated by substituting Equation (30) into Equation (29).

We also carried out a series of Gillespie simulations [29] of model 2 for several values of $\frac{\Delta g}{k_B T}$. Fig. 2 shows the position versus time for three example simulations, each carried out for a different value of $\frac{\Delta g}{k_B T}$. For each of these LEFs, the gray traces represent the positions of the junctions between nucleosomal DNA and naked DNA, the blue trace represents the position of the remodeling complex, and the red trace represents the position of the SMC complex. The mean position of the bottom LEF, which corresponds to $\frac{\Delta g}{k_B T} = 18$, remains essentially fixed over the period of the simulation, implying a very small LEF velocity. In addition, in this case, the remodeler and the SMC remain next to each other throughout the trajectory, implying a very small value of P_3. By contrast, the mean position of the middle LEF ($\frac{\Delta g}{k_B T} = 8$) increases more-or-less linearly in time with the remodeler and the SMC both stepping forward and frequently moving out of contact. Thus, in this case, the LEF shows a significant velocity and an intermediate value of P_3. Finally, although the velocity of the top LEF

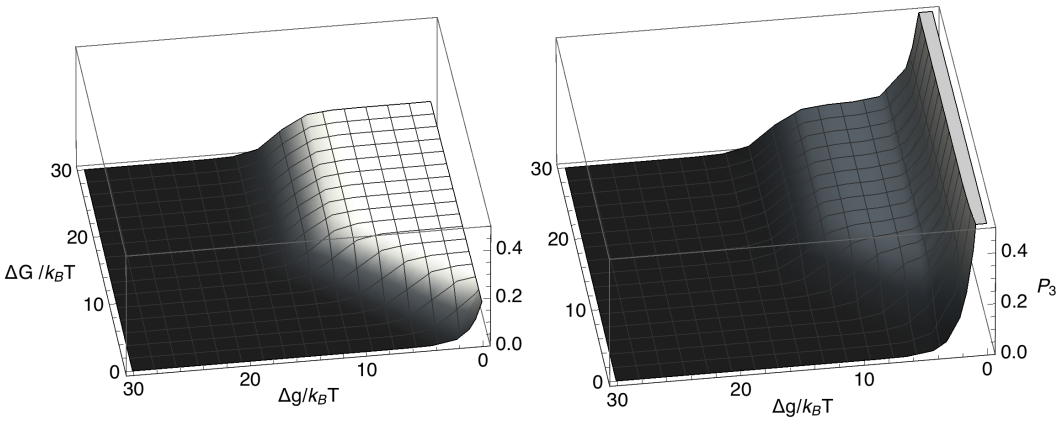

FIG. 3. Probability, P_3, that the remodeling complex and the SMC complex involved in a composite LEF are not adjacent to each other, plotted versus $\Delta G/(k_BT)$ and $\Delta g/(k_BT)$, according to model 1 [Equation (14)] for $k_+ = 0.05$ per time step, $k_- = 5.0 \times 10^{-7}$ per time step, $m_+ = 0.3$ per time step, and $m_- = 0.0003$ per time step (left) or $m_- = 0.3$ per time step (right).

($\frac{\Delta g}{k_BT} = 0.5$) is very similar to that of the middle LEF, the top LEF shows many fewer remodeler-SMC contacts than the middle LEF, corresponding to a significantly larger value of P_3. The cyan, green, and magenta lines in Fig. 2 have slopes given by the corresponding model-2 composite LEF velocities, – calculated by substituting Equation (30) into Equation (29) – revealing good agreement between theory and simulation.

C. Model 3

Model 3 supposes that the probability of complete nucleosome unbinding into solution is negligible, but that there exists a "remodeled" configuration, in which the nucleosome is both associated with the remodeler and also sufficiently displaced to allow the remodeler to step forward (bottom row of Fig. 1). In this model, we interpret Δg to be the free energy of the remodeled configuration. For simplicity, we also assume a separation of time scales with remodeling occurring much faster than translocation. Then, the probability that the remodeling complex and junction 1 are not next to each other is

$$P_1 = \frac{1}{1 + e^{\frac{\Delta g - \Delta G}{k_BT}}} = \frac{1}{1 + (1 + e^{\frac{\Delta g}{k_BT}})e^{\frac{\Delta G}{k_BT}} - e^{\frac{\Delta G}{k_BT}}}, \quad (31)$$

while the probability that the SMC and junction 2 are not next to each other is

$$P_2 = \frac{1}{1 + e^{\frac{\Delta g}{k_BT}}}. \quad (32)$$

Equations (31) and (32) replace model 1's Equations (1) and (2), respectively. However, Equations (9) and (12) are unchanged for model 3. It is apparent therefore that we may write down the model-3 results for P_3 and v by replacing $e^{\frac{\Delta g}{k_BT}}$ in corresponding results for model 1 by $e^{\frac{\Delta g}{k_BT}} + 1$. Thus, for model 3, we find

$$P_3 = \frac{\frac{k_+}{k_- m_+} + (\frac{1 - e^{-\frac{\Delta G}{k_BT}}}{1 + e^{\frac{\Delta g}{k_BT}}} + e^{-\frac{\Delta G}{k_BT}})\frac{m_-}{k_- m_+}}{\frac{1 + e^{\frac{\Delta g}{k_BT}}}{m_+} + \frac{1 + e^{\frac{\Delta g - \Delta G}{k_BT}}}{k_-}} \quad (33)$$

and

$$v = b \frac{\frac{k_+}{k_-} - \frac{m_-}{m_+}}{\frac{1 + e^{\frac{\Delta g}{k_BT}}}{m_+} + \frac{1 + e^{\frac{\Delta g - \Delta G}{k_BT}}}{k_-}}. \quad (34)$$

III. DISCUSSION

To realize a composite LEF, junction 1 and the remodeler, on the one hand, must not outrun the SMC and junction 2, on the other. This requirement may be expressed mathematically by insisting that the probability, P_3, that the remodeling complex and the SMC are not next to each other, must be less than 1. Otherwise, for $P_3 = 1$, the remodeler and SMC do not come into contact, and we may infer that the remodeler has outpaced the SMC. Fig. 3 plots P_3, according to model 1, as a function of $\frac{\Delta G}{k_BT}$ and $\frac{\Delta g}{k_BT}$. For the parameter values, used in the left-hand panel, we see that P_3 is everywhere less than 1, consistent with the existence of a composite LEF throughout the region illustrated. In fact, P_3 takes on a relatively large plateau value for

$$\Delta G > \Delta g \quad (35)$$

and

$$m_+ > e^{\frac{\Delta g}{k_BT}} k_-. \quad (36)$$

Elsewhere, P_3 is small.

For the parameter values used in the right-hand panel of Fig. 3, however, although P_3 shows a similar plateau at intermediate values of $\frac{\Delta g}{k_B T}$, as $\frac{\Delta g}{k_B T}$ decreases to near zero, P_3 increases rapidly to unity, and according to Equation (14), would unphysically exceed unity for small enough $\frac{\Delta g}{k_B T}$. This circumstance arises when even $P_3 = 1$ is not sufficient to satisfy $v_R = v_S$. When the remodeling complex and junction 1 outrun the SMC and junction 2 – *i.e.* when $v_R > v_S$ – the premise of a composite LEF, upon which Equations (14) and (15) are based, can no longer hold. Thus, to achieve a composite LEF, we must have that $v_R \leq v_S$ for $P_3 = 1$. This condition requires that the model parameter values must satisfy

$$m_+ - m_- e^{-\frac{\Delta g}{k_B T}} > \frac{k_+ - k_- e^{\frac{\Delta g}{k_B T}}}{1 + e^{\frac{\Delta g - \Delta G}{k_B T}} - e^{-\frac{\Delta G}{k_B T}}}. \qquad (37)$$

This condition is violated at small Δg for the parameters used in the right-hand panel of Fig. 3. For $k_+ \gg k_- e^{\frac{\Delta g}{k_B T}}$ and $m_+ \gg m_- e^{-\frac{\Delta g}{k_B T}}$, the condition for a composite LEF to exist becomes simply $m_+ > k_+$, namely the forward stepping rate of the SMC on naked DNA should be larger than the forward stepping rate of the remodeler on naked DNA.

To further elucidate the composite LEF's behavior as P_3 increases, we turned to Gillespie simulations of the sort illustrated in Fig. 2. The points in Fig. 4 show the simulated results for both P_3 itself (top panel) and the remodeler-SMC separation (bottom panel), plotted versus $\frac{\Delta g}{k_B T}$. The solid line in the top panel corresponds to Equation (30), demonstrating excellent quantitative agreement between theory and simulation for P_3. For the parameters of Fig. 4, as $\frac{\Delta g}{k_B T}$ decreases below about 3, P_3 increases from its plateau value, eventually reaching unity at $\frac{\Delta g}{k_B T} \simeq 0.2$. Thus, in this case, for $\frac{\Delta g}{k_B T} < 0.2$, a composite LEF does not exist.

It is apparent from the bottom panel of Fig. 4, that the remodeler-SMC separation matches P_3 for $\frac{\Delta g}{k_B T} \geq 3$. This result obtains because, for $\frac{\Delta g}{k_B T} \geq 3$, the overwhelmingly prevalent remodeler-SMC separations are 0 and 1, so that the calculation of P_3 and the calculation of the mean remodeler-SMC separation are effectively the same calculation in this regime. However, as $\frac{\Delta g}{k_B T}$ decreases below 3, the mean remodeler-SMC separation rapidly increases beyond P_3, as larger remodeler-SMC separations than 1 become prevalent, as may seen for the top LEF in Fig. 2, which corresponds to $\frac{\Delta g}{k_B T} = 0.5$. The mean remodeler-SMC separation reaches 1 for $\frac{\Delta g}{k_B T} \lesssim 1.3$ and rapidly increases as $\frac{\Delta g}{k_B T}$ decreases further.

A key assumption of our theory is that displaced nucleosomes rebind only at junctions between nucleosomal DNA and naked DNA. However, when the model predicts a relatively large region of naked DNA between the remodeler and the SMC, into which a nucleosome could easily fit, this assumption seems likely to be inappropriate and the model no longer self-consistent, in turn suggesting that the condition specified by Equation (37)

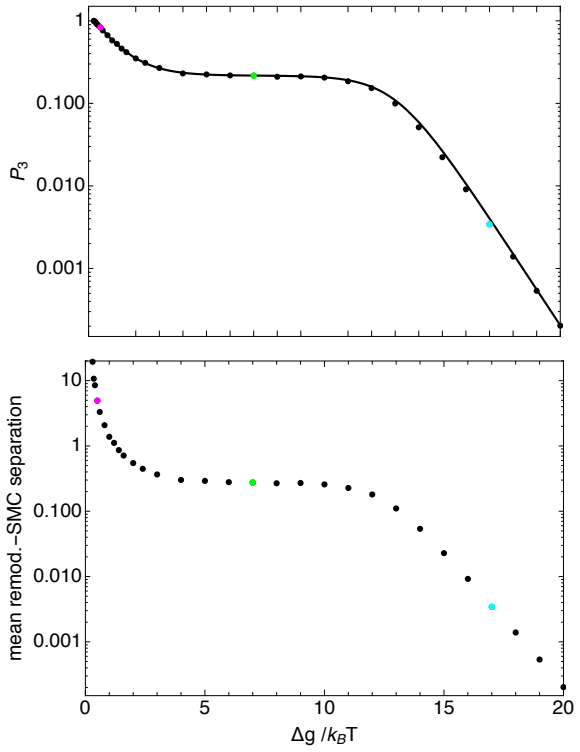

FIG. 4. The probability, P_3, that the remodeler and SMC are not next to each other (top) and the mean remodeler-SMC separation (bottom), plotted versus nucleosome binding energy, $\frac{\Delta g}{k_B T}$. The circles correspond to results determined from model-2 Gillepsie simulations, each containing 2^{20} transitions. The solid line corresponds to Equation (30). The parameter values used were: $k_+ = 0.05$ per time step, $k_- = 5 \times 10^{-7}$ per time step, $m_+ = m_- = 0.3$ per time step, $\Delta G = 18.0 k_B T$, $\alpha = 1$ per time step, and $\beta = \alpha e^{-\frac{\Delta g}{k_B T}}$. These parameters correspond to those for Fig. 2. The cyan, green, and magenta points at $\frac{\Delta g}{k_B T} = 0.5$, 8.0, and 18, respectively, correspond to the bottom, middle, and top traces of Fig. 2.

may be too permissive. However, further investigation of this question lies beyond the simple model described here.

Fig. 5 plots the model-1 LEF velocity, corresponding to the probabilities displayed in Fig. 3, showing that v achieves a relatively large plateau value when the conditions,

$$\Delta G > \Delta g \qquad (38)$$

and

$$m_+ > e^{\frac{\Delta g}{k_B T}} k_-, \qquad (39)$$

are both satisfied. Equation (38) informs us that to achieve rapid composite LEF translocation, a large repulsive nucleosome-remodeling complex interaction (ΔG) is necessary, that overcomes the nucleosome binding free energy (Δg). We might have expected that rapid composite LEF translocation would also require that the rate

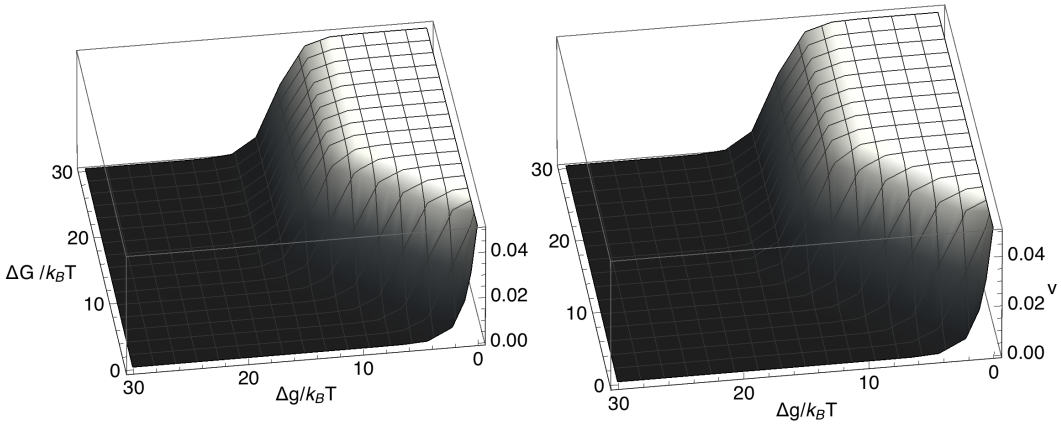

FIG. 5. Mean velocity, v, of a composite LEF plotted versus $\frac{\Delta G}{k_B T}$ and $\frac{\Delta g}{k_B T}$, according to Equation (15) for $k_+ = 0.05$ per time step, $k_- = 5.0 \times 10^{-7}$ per time step, $m_+ = 0.3$ per time step, and $m_- = 0.0003$ per time step (left) or $m_- = 0.3$ per time step (right).

at which the SMC complex steps forward into a gap between the SMC complex and the remodeling complex must exceed the rate at which the remodeling complex steps backwards into that same gap, which is $e^{\frac{\Delta G}{k_B T}} k_-$, *i.e.* we might have expected that $m_+ > e^{\frac{\Delta G}{k_B T}} k_-$. However, because of Equation (38), Equation (39) is actually a weaker condition on m_+ than this expectation.

Fig. 6 illustrates the model 1 diffusivities of the remodeling complex and the SMC complex. Each diffusivity specifies the corresponding factor's positional fluctuations, about the mean displacement, determined by the velocity. The diffusivities also show relatively large plateau values when Equations (38) and (39) are satisfied. Surprisingly, the diffusivity of the remodeling complex also shows a second plateau with an even higher plateau value for $\Delta G > \Delta g$ and $m_+ < e^{\frac{\Delta g}{k_B T}} k_-$, where the corresponding composite LEF velocity is small.

When all of Equations (37), (38) and (39) are simultaneously satisfied, the plateau values of the the probability that the remodeling complex and the SMC complex are not next to each other, the LEF velocity, and the two diffusivities are given approximately by

$$P_3 \simeq \frac{k_+}{m_+}, \tag{40}$$

$$v \simeq b k_- (\frac{k_+}{k_-} - \frac{m_-}{m_+}), \tag{41}$$

$$D_R \simeq \frac{1}{2} b^2 k_- (\frac{k_+}{k_-} + \frac{m_-}{m_+}), \tag{42}$$

and

$$D_S \simeq \frac{1}{2} b^2 k_+, \tag{43}$$

respectively. The plateau value of the composite LEF's loop extrusion velocity is independent of Δg. This result is possible (although not required – see below) because a loop extrusion step does not lead to a net change in the nucleosome configuration.

Fig. 5 shows that the LEF velocity is inevitably small for small ΔG. For $\Delta G = 0$, corresponding to solely hard-core repulsions between the remodeler and a nucleosome – what could be termed a "passive" composite LEF, in analogy to the passive helicase, discussed for example in Ref. [28] – Equation (15) becomes

$$v = b e^{-\frac{\Delta g}{k_B T}} \frac{\frac{k_+}{k_-} - \frac{m_-}{m_+}}{\frac{1}{m_+} + \frac{1}{k_-}}. \tag{44}$$

In this case, the composite LEF velocity decreases exponentially with the free energy of nucleosome unbinding, Δg. Since Δg is several tens of $k_B T$, we do not expect this limit to be feasible for effective loop extrusion. Although Equation (44) corresponds to $f = 0$ and $\Delta G = 0$, it may be shown that it also gives the LEF velocity for $f = 1$ in the large-ΔG limit. This is because for $f = 1$, large ΔG effectively creates a hard wall for the remodeler, albeit located one step away from the nucleosome, recapitulating the situation considered for $f = 0$ and $\Delta G = 0$.

In comparison to Equation (15), the velocity of a lonely remodeling complex, translocating on nucleosomal DNA, unaccompanied by an SMC complex, is

$$\begin{aligned} v_R &= b k_+ P_1 - b k_- e^{-\frac{\Delta g}{k_B T}} (P_1 + (1 - P_1) e^{\frac{\Delta G}{k_B T}}) \\ &= b \frac{k_+ - k_-}{1 + e^{\frac{\Delta g - \Delta G}{k_B T}} - e^{-\frac{\Delta G}{k_B T}}}, \end{aligned} \tag{45}$$

which may be straightforwardly obtained from Equation (10) by replacing P_3 with $e^{-\frac{\Delta g}{k_B T}}$, which is the probability that there is a gap between the remodeler and junction 2. The velocity of such a lonely remodeling complex

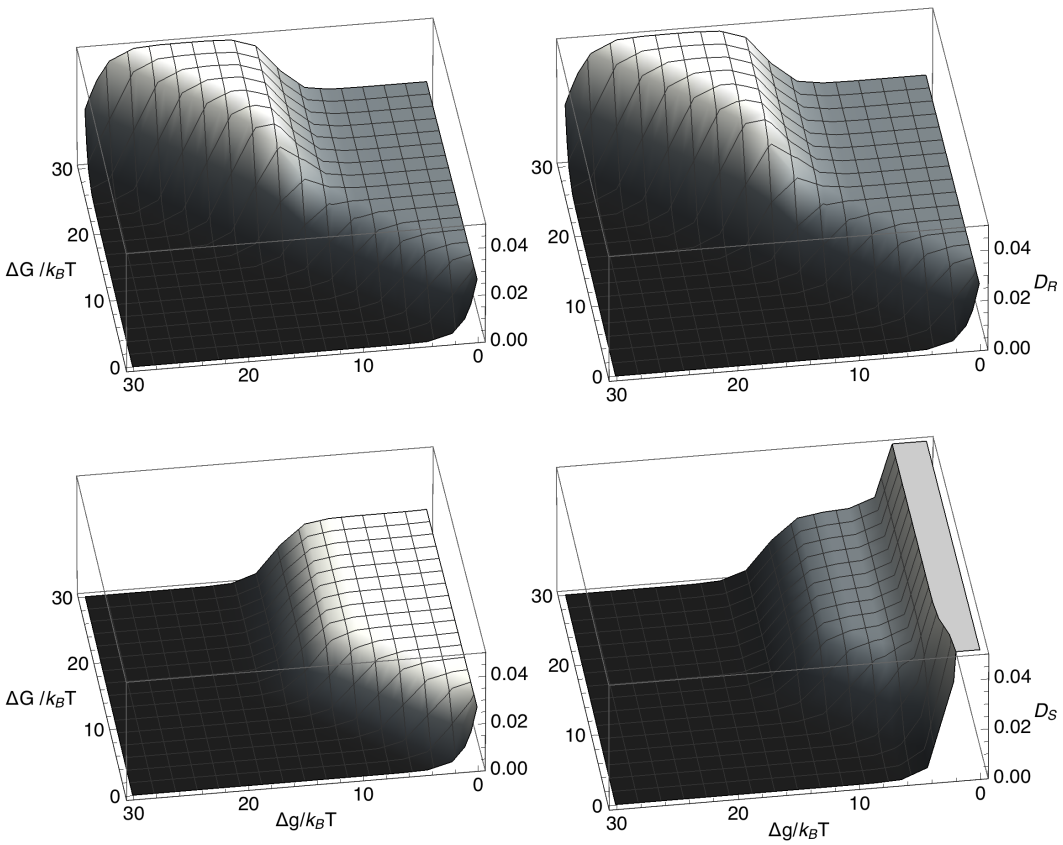

FIG. 6. Diffusivities, D_R (top row) and D_S (bottom row) of the remodeling complex and the SMC complex, respectively, plotted versus $\Delta G/(k_BT)$ and $\Delta g/(k_BT)$, according to model 1 [Equations (16) and (17)] for $k_+ = 0.05$ per time step, $k_- = 5.0 \times 10^{-7}$ per time step, $m_+ = 0.3$ per time step, and $m_- = 0.0003$ per time step (left column) or $m_- = 0.3$ per time step (right column).

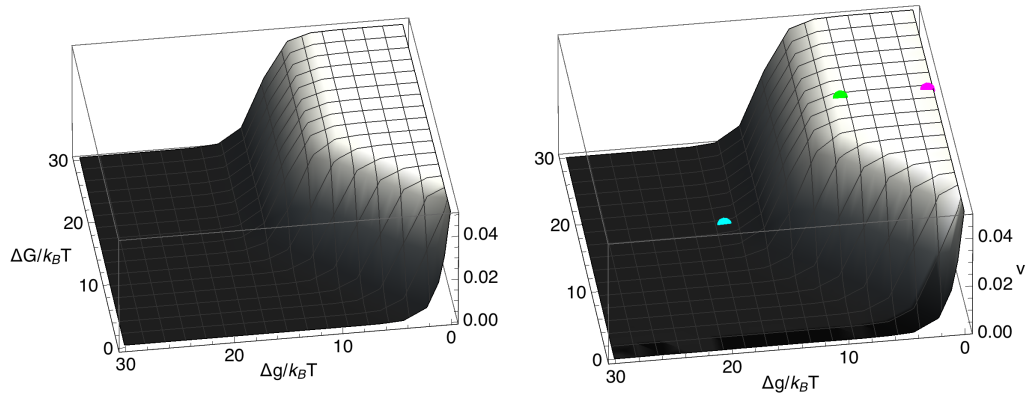

FIG. 7. Mean velocity, v, of a composite LEF, plotted versus $\Delta G/(k_BT)$ and $\Delta g/(k_BT)$ for model 1 (left) and model 2 (right) for $k_+ = 0.05$ per time step, $k_- = 5.0 \times 10^{-7}$ per time step, $m_+ = 0.3$ per time step, $m_- = 0.0003$ per time step, and (for model 2) $\alpha = 1$ per time step. The cyan, green, and magenta points on the model-2 curve correspond to the theoretical mean velocities of the composite LEFs whose positions versus time are shown in Fig. 2.

is relatively large for $\Delta G > \Delta g$ and is small otherwise. Thus, as seems intuitive, for efficient remodeler translocation on chromatin the remodeler-nucleosome repulsive free energy, ΔG must exceed the free energy required for nucleosome unbinding, Δg. In the large-ΔG limit, the remodeler velocity realizes a plateau value of

$$v_R = b(k_+ - k_-) = bk_-(\frac{k_+}{k_-} - 1), \tag{46}$$

so that the plateau velocity of a composite LEF exceeds (is less than) [equals] that of a lonely remodeling complex for $m_+ > m_-$ ($m_+ < m_-$) [$m_+ = m_-$].

We can also straightforwardly calculate the velocity of the SMC complex on nucleosomal DNA in the absence of the remodeling complex with the result that

$$v_S = be^{-\frac{\Delta g}{k_B T}}(m_+ - m_-). \tag{47}$$

Equation (47) informs us that, on nucleosomal DNA, the velocity of loop extrusion by an isolated SMC complex, which by assumption does not have its own nucleosome remodeling activity, is suppressed by a factor $e^{-\frac{\Delta g}{k_B T}}$ compared to the velocity of its loop extrusion on nucleosome-free DNA, which is $b(m_+ - m_-)$. Since $e^{-\frac{\Delta g}{k_B T}}$ is tiny, the velocity of the SMC without the remodeling complex is correspondingly tiny, even for $m_+ \gg m_-$, emphasizing that the remodeling complex is essential for significant loop extrusion in the chromatin context.

Equation (15) informs us that the composite LEF's directionality depends only on $\frac{k_+}{k_-} - \frac{m_-}{m_+}$. Since we can expect that $\frac{k_+}{k_-} = e^{-\frac{\Delta G_R}{k_B T}}$ and $\frac{m_+}{m_-} = e^{-\frac{\Delta G_S}{k_B T}}$, where ΔG_R is the free energy change associated with the remodeling complex stepping forward and ΔG_S is the free energy change, associated with the SMC complex stepping forward, it is clear that the composite LEF proceeds forward, only provided $\Delta G_R + \Delta G_S < 0$. This outcome reflects the Second Law of Thermodynamics, expressed in the form that a chemical reaction proceeds forward only if the corresponding change in free energy is negative. In comparison, Equation (45) informs us that a lonely remodeling complex proceeds forwards if $\Delta G_R < 0$.

Shown in Fig. 7 is a comparison between the LEF velocity for model 2 and the LEF velocity for model 1. Model 2 reproduces both the region in the ΔG-Δg plane where the composite LEF velocity is large and the plateau value of the LEF velocity within that region [Equation (41)]. The cyan, green, and magenta points on the model-2 curve in Fig. 7 correspond to the free energy settings and theoretical mean velocities of the composite LEFs, whose simulated positions versus time are shown in Fig. 2. Both the top group of traces and the middle group of traces in Figure 2 fall within the plateau region of the velocity, which explains why their velocities are very similar. However, while the middle LEF does fall within the plateau region of P_3, the top composite LEF exhibits a significantly large value of P_3 and a correspondingly larger spatial extent.

The conceptually simplest versions of the composite LEF model (models 1 and 2) envision that the remodeler ejects a nucleosome from the DNA ahead of the remodeler, and that the nucleosome subsequently rebinds behind the SMC. Alternatively, model 3 hypothesizes an intermediate, "remodeled" state in which the displaced nucleosome remains associated with the LEF, eventually to relocate behind SMC. This picture is reminiscent of the scenario envisioned in Ref. [30], which demonstrated experimentally that RNA polymerase could pass a nucleosome without causing nucleosome dissociation. Nonetheless, for $e^{\frac{\Delta g}{k_B T}} \gg 1$, the predictions of all three models are indistinguishable. The interpretation of Δg is different for models 1 and 2, on the one hand, and model 3 on the other. For models 1 and 2, Δg is the nucleosome binding free energy, which is several tens of $k_B T$. For model 3, Δg is the free energy of the remodeled configuration, relative to the free energy of a bound nucleosome, which we may expect to be smaller than the free energy required to nucleosome unbinding (models 1 and 2). However, as noted above, the plateau value of the composite LEF's loop extrusion velocity is independent of Δg for all of the models.

Recent single-molecule measurements demonstrate that the motion of SMCs on DNA is blocked by sufficiently large DNA-bound proteins, including in particular RNA polymerase [20], which is one possible candidate remodeler component of a composite LEF. Thus, our premise that the SMC cannot pass the remodeler seems justifiable.

Other experiments have indicated that condensin takes steps on naked DNA, that are up to 600 bp in size [31, 32], using the free energy from two ATP hydrolysis events to do so [16]. By contrast, single base pair steps are involved in the remodeling activity of canonical chromatin remodelers, such as SWI/SNF, ISW1, and INO80 [33]. In addition, RNA polymerase, which also possesses remodeling activity, necessarily takes single-base pair steps. In these cases, a step involves hydrolysis of at least one ATP molecule. As a result, because the number of ATPs per base pair is far larger for a remodeler than for an SMC complex, we should expect that $\Delta G_R \gg \Delta G_S$, and therefore that $\frac{k_+}{k_-} \gg \frac{m_-}{m_+}$. Thus, a prediction of our model is that the composite LEF velocity is essentially determined by the velocity of its component remodeler and not by the velocity of the SMC complex on nucleosome-free DNA.

It is interesting to ask whether there is any experimental support for this prediction. Ref. [34] estimates that the LEF velocity on chromatin is a few tens of nm s^{-1}. Specifically, by comparing the results of Hi-C measurements to loop-extrusion simulations, researchers have estimated that in higher eukaryotes LEFs give rise to 100 kilobase loops on average [10, 35–37]. Such loops develop during the residence time of an SMC on chromatin, which is about 1000 s [38–42]. On the basis of these two estimates, Ref. [34] infers that the velocity of loop extrusion on chromatin is about 30 nm s^{-1}. Remarkably, this estimated loop extrusion velocity is many

times smaller than the velocity of loop extrusion on naked DNA *in vitro*, which is 500 nm s^{-1} [16], but is comparable to the velocity *in vivo* of RNA polymerase, which is 10-30 nm s^{-1} [43]. The agreement between the velocity of remodeler-candidate RNA polymerase and the estimated velocity of loop extrusion *in vivo* seems consistent with the composite LEF model.

IV. CONCLUSIONS

A key result of this paper is that even if nucleosomes block SMC translocation, efficient loop extrusion remains possible on chromatinized DNA via a LEF, that is a composite entity involving a remodeler and nucleosomes, as well as an SMC complex. Thus, the possibility that nucleosomes may block SMC translocation and loop extrusion on chromatin is *not* a reason to rule out the loop extrusion factor model of genome organization.

We have shown that, for a wide range of possible parameter values, such a composite LEF exists as a more-or-less compact entity with all its component parts in close proximity to each other, and can give rise to loop extrusion with a velocity, that is comparable to the remodeler's translocation velocity on chromatin, but is much larger than the velocity of a SMC complex that is blocked by nucleosomes. Although we have focused on one-sided loop extrusion, two-sided loop extrusion simply requires two remodelers, one for each chromatin strand threading the SMC.

The composite LEF model is agnostic concerning whether the SMC complex shows ATP-dependent translocase activity ($m_+ \neq m_-$) or diffuses ($m_+ = m_-$) on naked DNA. However, Equation 37 specifies the condition for a composite LEF to exist defined by the SMC and the remodeler being in close proximity, while efficient chromatin loop extrusion requires repulsion between the remodeler and the junction between nucleosomal DNA and naked DNA, that is large compared to the nucleosome binding free energy (models 1 and 2) or the remodeled configuration free energy (model 3): $\Delta G > \Delta g$. An additional condition necessary for efficient loop extrusion is $m_+ > e^{\frac{\Delta g}{k_B T}} k_-$. Finally, we remark that the composite LEF model, described in this paper, is quite distinct from the models of Refs. [44, 45], which propose loop extrusion occurs without the involvement of a translocase.

ACKNOWLEDGMENTS

This research was supported by NSF CMMI 1634988 and NSF EFRI CEE award EFMA-1830904. M. L. P. B. was supported by NIH T32EB019941 and the NSF GRFP.

[1] https://bionumbers.hms.harvard.edu.
[2] J. Dekker, K. Rippe, M. Dekker, and N. Kleckner, Capturing chromosome conformation, Science **295**, 1306 (2002).
[3] J. R. Dixon, S. Selvaraj, F. Yue, A. Kim, Y. Li, Y. Shen, M. Hu, J. S. Liu, and B. Ren, Topological domains in mammalian genomes identified by analysis of chromatin interactions, Nature **485**, 376 (2012).
[4] J. R. Dixon, D. U. Gorkin, and B. Ren, Chromatin domains: The unit of chromosome organization, Mol Cell **62**, 668 (2016).
[5] T. Sexton, E. Yaffe, E. Kenigsberg, F. Bantignies, B. Leblanc, M. Hoichman, H. Parrinello, A. Tanay, and G. Cavalli, Three-dimensional folding and functional organization principles of the Drosophila genome, Cell **148**, 458 (2012).
[6] T. Mizuguchi, G. Fudenberg, S. Mehta, J. M. Belton, N. Taneja, H. D. Folco, P. FitzGerald, J. Dekker, L. Mirny, J. Barrowman, and S. I. Grewal, Cohesin-dependent globules and heterochromatin shape 3D genome architecture in *:s. pombe*, Nature **516**, 432 (2014).
[7] J. Dekker, Two ways to fold the genome during the cell cycle: Insights obtained with chromosome conformation capture, Epigenetics Chromatin **7**, 25 (2014).
[8] E. Alipour and J. F. Marko, Self-organization of domain structures by DNA-loop-extruding enzymes, Nucleic acids research **40**, 11202 (2012).
[9] A. L. Sanborn, S. S. Rao, S. C. Huang, N. C. Durand, M. H. Huntley, A. I. Jewett, I. D. Bochkov, D. Chinnappan, A. Cutkosky, J. Li, K. P. Geeting, A. Gnirke, A. Melnikov, D. McKenna, E. K. Stamenova, E. S. Lander, and E. L. Aiden, Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes, Proc Natl Acad Sci U S A **112**, E6456 (2015).
[10] G. Fudenberg, M. Imakaev, C. Lu, A. Goloborodko, N. Abdennur, and L. A. Mirny, Formation of chromosomal domains by loop extrusion, Cell Reports **15**, 2038 (2016).
[11] J. Nuebler, G. Fudenberg, M. Imakaev, N. Abdennur, and L. A. Mirny, Chromatin organization by an interplay of loop extrusion and compartmental segregation, Proc Natl Acad Sci U S A **115**, E6697 (2018).
[12] A. Goloborodko, J. F. Marko, and L. A. Mirny, Chromosome compaction by active loop extrusion, Biophysical journal **110**, 2162 (2016).
[13] A. Goloborodko, M. V. Imakaev, J. F. Marko, and L. Mirny, Compaction and segregation of sister chromatids via active loop extrusion, eLife **5**, e14864 (2016).
[14] Y. Murayama and F. Uhlmann, Biochemical reconstitution of topological DNA binding by the cohesin ring, Nature **505**, 367 (2014).
[15] A. Khoury, J. Achinger-Kawecka, S. A. Bert, G. C. Smith, H. J. French, P.-L. Luu, T. J. Peters, Q. Du, A. J. Parry, F. Valdes-Mora, P. C. Taberlay, C. Stirzaker, A. L. Statham, and S. J. Clark, Constitutively bound CTCF

sites maintain 3D chromatin architecture and long-range epigenetically regulated domains, Nature Communications **11**, 54 (2020).

[16] M. Ganji, I. A. Shaltiel, S. Bisht, E. Kim, A. Kalichava, C. H. Haering, and C. Dekker, Real-time imaging of DNA loop extrusion by condensin, Science **360**, 102 (2018).

[17] Y. Kim, Z. Shi, H. Zhang, I. J. Finkelstein, and H. Yu, Human cohesin compacts DNA by loop extrusion, Science **366**, 1345 (2019).

[18] M. Kong, E. E. Cutts, D. Pan, F. Beuron, T. Kaliyappan, C. Xue, E. P. Morris, A. Musacchio, A. Vannini, and E. C. Greene, Human condensin I and II drive extensive ATP-dependent compaction of nucleosome-bound DNA, Molecular Cell **79**, 99 (2020).

[19] J. Stigler, G. O. Camdere, D. E. Koshland, and E. C. Greene, Single-molecule imaging reveals a collapsed conformational state for DNA-bound cohesin, Cell Rep **15**, 988 (2016).

[20] I. F. Davidson, D. Goetz, M. P. Zaczek, M. I. Molodtsov, P. J. H. in 't Veld, F. Weissmann, G. Litos, D. A. Cisneros, M. Ocampo-Hafalla, R. Ladurner, F. Uhlmann, A. Vaziri, and J.-M. Peters, Rapid movement and transcriptional re-localization of human cohesin on DNA, EMBO J **35**, 2671 (2016).

[21] R. N. Dubey and M. R. Gartenberg, A tDNA establishes cohesion of a neighboring silent chromatin domain, Genes Dev **21**, 2150 (2007).

[22] E. F. Glynn, P. C. Megee, H. G. Yu, C. Mistrot, E. Unal, D. E. Koshland, J. L. DeRisi, and J. L. Gerton, Genome-wide mapping of the cohesin complex in the yeast *saccharomyces cerevisiae*, PLoS Biol **2**, E259 (2004).

[23] A. Lengronne, Y. Katou, S. Mori, S. Yokobayashi, G. P. Kelly, T. Itoh, Y. Watanabe, K. Shirahige, and F. Uhlmann, Cohesin relocation from sites of chromosomal loading to places of convergent transcription, Nature **430**, 573 (2004).

[24] C. K. Schmidt, N. Brookes, and F. Uhlmann, Conserved features of cohesin binding along fission yeast chromosomes, Genome Biol **10**, R52 (2009).

[25] S. Golfier, T. Quail, H. Kimura, and J. Brugués, Cohesin and condensin extrude loops in a cell-cycle dependent manner, eLife **9**, e53885 (2020).

[26] C. S. Peskin, G. M. Odell, and G. F. Oster, Cellular motions and thermal fluctuations: the Brownian ratchet, Biophys J. **65**, 316 (1993).

[27] M. D. Betterton and F. Jülicher, A motor that makes its own track: Helicase unwinding of DNA, Physical Review Letters **91**, 258103 (2003).

[28] M. D. Betterton and F. Jülicher, Opening of nucleic-acid double strands by helicases: Active versus passive opening, Phys. Rev. E **71**, 011904 (2005).

[29] D. T. Gillespie, Exact stochastic simulation of coupled chemical reactions, The Journal of Physical Chemistry **81**, 2340 (1977).

[30] C. Hodges, L. Bintu, L. Lubkowska, M. Kashlev, and C. Bustamante, Nucleosomal fluctuations govern the transcription dynamics of RNA polymerase II, Science **325**, 626 (2009).

[31] J.-K. Ryu, S.-H. Rah, R. Janissen, J. W. J. Kerssemakers, and C. Dekker, Resolving the step size in condensin-driven dna loop extrusion identifies atp binding as the step-generating process, bioRxiv 10.1101/2020.11.04.368506 (2020).

[32] T. R. Strick, T. Kawaguchi, and T. Hirano, Real-time detection of single-molecule dna compaction by condensin i, Current Biology **14**, 874 (2004).

[33] L. Yan and Z. Chen, A unifying mechanism of dna translocation underlying chromatin remodeling, Trends in Biochemical Sciences **45**, 217 (2020).

[34] E. J. Banigan and L. A. Mirny, Loop extrusion: theory meets single-molecule experiments, Current Opinion in Cell Biology **64**, 124 (2020), cell Nucleus.

[35] J. Gassler, H. Brandao, M. Imakaev, I. Flyamer, S. Ladstatter, W. Bickmore, J. Peters, L. Mirny, and K. Tachibana, A mechanism of cohesin-dependent loop extrusion organizes zygotic genome architecture, EMBO J **36**, 3600 (2017).

[36] J. Gibcus, K. Samejima, A. Goloborodko, I. Samejima, N. Naumova, J. Nuebler, M. Kanemaki, L. Xie, J. Paulson, W. Earnshaw, L. Mirny, and J. Dekker, A pathway for mitotic chromosome formation, Science **359** (2018).

[37] E. Banigan, A. van den Berg, H. Brandao, J. Marko, and L. Mirny, Chromosome organization by one-sided and two-sided loop extrusion, eLife **9**, e53558 (2020).

[38] G. Wutz, C. Varnai, K. Nagasaka, D. A. Cisneros, R. R. Stocsits, W. Tang, S. Schoenfelder, G. Jessberger, M. Muhar, M. J. Hossain, N. Walther, B. Koch, M. Kueblbeck, J. Ellenberg, J. Zuber, P. Fraser, and J. M. Peters, Topologically associating domains and chromatin loops depend on cohesin and are regulated by CTCF, WAPL, and PDS5 proteins, EMBO J **36**, 3573 (2017).

[39] S. Kueng, B. Hegemann, B. Peters, J. Lipp, A. Schleiffer, K. Mechtler, and J.-M. Peters, Wapl controls the dynamic association of cohesin with chromatin, Cell **127**, 955 (2006).

[40] A. McNairn and J. Gerton, Intersection of ChIP and FLIP, genomic methods to study the dynamics of the cohesin proteins, Chromosome Res. **17**, 155 (2009).

[41] A. Tedeschi, G. Wutz, S. Huet, M. Jaritz, A. Wuensche, E. Schirghuber, I. Davidson, W. Tang, D. Cisneros, V. Bhaskara, T. Nishiyama, A. Vaziri, A. Wutz, J. Ellenberg, and J.-M. Peters, Wapl is an essential regulator of chromatin structure and chromosome segregation, Nature **501**, 564 (2013).

[42] A. Hansen, I. Pustova, C. Cattoglio, R. Tjian, and X. Darzacq, CTCF and cohesin regulate chromatin loop stability with distinct dynamics, Elife **6**, 10.7554/eLife.25776 (2017).

[43] M. B. Ardehali and J. T. Lis, Tracking rates of transcription and splicing *in vivo*, Nat. Struct. Mol. Biol. **16**, 1123 (2009).

[44] C. A. Brackley, J. Johnson, D. Michieletto, A. N. Morozov, M. Nicodemi, P. R. Cook, and D. Marenduzzo, Nonequilibrium chromosome looping via molecular slip links, Physical Review Letters **119**, 138101 (2017).

[45] A. Maji, R. Padinhateeri, and M. K. Mitra, The accidental ally: Nucleosome barriers can accelerate cohesin-mediated loop formation in chromatin, Biophys. J. **119**, 2316 (2020).