DNA-segment-capture model for loop extrusion by structural maintenance of chromosome (SMC) protein complexes

John F. Marko

Department of Molecular Biosciences and Department of Physics and Astronomy, Northwestern University, Evanston IL 60208 USA

Paolo De Los Rios

Laboratory of Statistical Biophysics, Institute of Physics, School of Basic Sciences and Institute of Bioengineering, School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne - EPFL, 1015 Lausanne, Switzerland

Alessandro Barducci

Center de Biochimie Structurale, INSERM, CNRS, Université de Montpellier, 34090 Montpellier, France

Stephan Gruber

Department de microbiologie fondamentale, Université de Lausanne, 1015 Lausanne, Switzerland

Abstract

Cells possess remarkable control of the folding and entanglement topology of long and flexible chromosomal DNA molecules. It is thought that structural maintenance of chromosome (SMC) protein complexes play a crucial role in this, by organizing long DNAs into series of loops. Recent experimental data suggest that SMC complexes are able to translocate on DNA, as well as pull out lengths of DNA via a "loop extrusion" process. We describe a Brownian loop-capture-ratchet model for translocation and loop extrusion based on known structural, catalytic, and DNA-binding properties of the *Bacillus subtilis* SMC complex. Our model provides an example of a new class of molecular motor where large conformational fluctuations of the motor 'track' - in this case DNA - are involved in the basic translocation process. Quantitative analysis of our model leads to a series of predictions for the motor properties of SMC complexes, most strikingly a strong dependence of SMC translocation velocity on tension in the DNA track that it is moving along, with "stalling" occuring at subpiconewton tensions. A closely related hallmark of the loop-capture mechanism is a strong dependence of translocation step size on DNA tension. We also discuss how the same mechanism might be used by other structurally related SMC complexes (*E. coli* MukBEF and eukaryote condensin, cohesin and SMC5/6) to organize genomic DNA.

I. INTRODUCTION

A fundamental problem faced by all living cells is dealing with the fantastic length of their chromosomal DNAs - millimeters in bacteria, and centimeters to meters in many eukaryote cells. Given the ≈ 300 base pair (bp) statistical segment length of double-stranded DNA (dsDNA), the $\approx 5 \times 10^6$ bp chromosomes of bacteria and the $\approx 10^8$ bp chromosomes of mammalian cells are in the realm of extremely long and flexible polymers. Left to their own devices, the multiple chromosomal DNAs of these lengths found in bacteria and in eukaryote cell nuclei should become highly entangled with one another [1–3]. Nevertheless, cells are able to completely separate their chromosomal DNAs from one another, and to keep them from being entangled with one another.

Two enzymatic machines are believed to play an essential role in the topological simplification of chromosomal DNAs *in vivo*. The first are type II topoisomerases (topos), which are enzymes which pass one dsDNA through a catalytically generated break in a second dsDNA [4]. Following strand passage, type II topos reseal the dsDNA break and then release the DNA strands, effecting a change of crossing sign that is essential to changing knotting or linking topology [5] (Fig. 1a). The key enzyme of this type in bacteria is called topo IV, while in eukaryote cells it is topo IIα. Type II topos are ATPases, and use stored energy to complete their reaction cycle. Although type II topos are known to have the capacity to channel energy released by use of ATP into simplification of entanglement topology of DNAs to levels below that expected in thermal equilibrium [6–8], this effect is insufficient by itself to eliminate entanglement of whole chromosomes in cells.

A second class of enzymatic machine which appears essential to topological simplification of whole chromosomes are Structural Maintenance of Chromosome (SMC) complexes, which are also ATP-consuming molecular machines (Fig. 1b) [9–12]. SMC complexes are based on $\approx$50-nanometer (nm) long coiled-coil protein domains that close into ring-shaped structures (Fig. 2). An array of experimental data point to the capacity of SMC complexes to translocate along DNA in a directed fashion, as well as to mediate DNA "loop extrusion" processes, whereby an initially small DNA loop is increased in size processively. SMC complexes are thought to be involved in a number of DNA-organizing processes, including tethering of the

two halves of bacterial chromosomes together [13–17], DNA-sequence-nonspecific compaction of chromosomes via crowding together of DNA loops [3, 18, 19], and also the bringing together of specific signal sequences spaced by distances of 10^5 to 10^6 bp along DNA molecules in specific orientations [20, 21]. Recent single-molecule experiments on yeast condensin have directly observed translocation [22] and loop-extrusion [23] functions.

At their core, these functions involve organization of long dsDNAs into series of loop structures, and plausibly facilitate detection of DNA connectivity and control of DNA topology in a way that appears impossible without a directed translocation-based mechanism. Impressively, SMC complexes appear to be an ancient class of enzymes, with structurally similar complexes occuring in all kingdoms of life. Key SMC complexes under current study include the bsSMC complex in the bacterium *Bacillus subtilus*, the MukBEF complex found in the bacterium *E. coli*, and the condensin, cohesin and SMC5/6 complexes found in eukaryote cells.

A major puzzle regarding SMC complexes has been exactly how they are able to translocate and loop-extrude DNA contour, given that they are structurally unrelated to known DNA-based protein machines, e.g., RNA polymerase, DNA polymerase and DNA helicases. Structural and enzymatic analyses of the bsSMC complex have led to a model whereby small ($\approx$ 600 bp) loops of DNA are captured and incorporated together to gradually extrude loops [24]. Here, we analyze this model from a quantitative theoretical biophysics point of view. We emphasize that although we focus on bsSMC, much of our modeling appears to be applicable to eukaryote condensin, which has been the focus of a great deal of recent attention.

In Section II we begin by describing the sequence of conformational states thought to occur for bsSMC during its ATP binding, hydrolysis, and product release cycle, and then we describe how DNA conformational changes can be coupled to these protein conformational states so as to accomplish translocation of a SMC along a dsDNA molecule. An essential ingredient of the model is the capture of thermally generated DNA loops, which leads to a strong dependence of the translocation rate on tension in the dsDNA 'track', illustrating its uniqueness as a molecular motor system where the flexibiity and deformation of the motor track plays an essential role. Our model is sufficiently biophysically based that we can make predictions about the transition rates between the different SMC/DNA states and therefore the cycling and translocation dynamics of the SMC/DNA combination.

In Section III we describe a few plausible schemes whereby DNA loops could be extruded by SMC complexes: in one, two SMC translocators are coupled together, but in the others we show how a single SMC complex may be able to generate processive DNA loop extrusion. Finally, in the Discussion section we review the key results of the paper, and predictions of the model for specific experiments.

II. MODEL FOR SMC TRANSLOCATION ALONG DNA

While we will focus on the specific case of bsSMC (where there is at present the best structural understanding), the model we now construct will be general enough to be applicable to all SMC complexes. Fig. 2 shows the 'body plan' of bsSMC: like other SMCs they consist of long coil-colied proteins (called SMC proteins) which dimerize at one end at "hinge" domains, and which have 'head' domains at their other ends which bind and hydrolyse ATP. The two heads are able to bind together when they bind ATP, and ATP hydrolysis releases this binding.

All SMC complexes also contain an additional protein, often called a "kleisin" subunit, which bridges the two heads using two distinct interfaces, breaking the symmetry of the SMC dimer (the two SMC proteins are a homodimer in the bacterial case, and a highly symmetric heterodimer in the eukaryote case). Various species of SMC have additional subunits which likely play key roles in DNA targeting and regulation of the complex, but for this paper we will focus on the interplay between the hinge-coiled-coil-head SMC proteins and the bridging protein.

A. Structural states of SMC complexes

SMC complexes can be found in three distinct nucleotide-binding states, either the "apo" state with no ATP bound, the ATP-bound state, or the ADP-bound state [24]. *In vivo*, ATP concentration is essentially held fixed, which indicates that these three states define a chemical reaction cycle: the apo state (Fig. 3a) will bind free ATPs (Fig. 3b), then hydrolyse them (Fig. 3c), and then release the hydrolysis products to return to the apo state (Fig. 3a). While there may be additional states that a particular SMC complex might be able to occupy, all SMC complexes are thought to go through this ATP binding/hydrolysis cycle in

vivo.

For the case of bsSMC, it is established that these three ATP-binding states correspond to different structural states of the enzyme. In the apo state, the long coiled-coil arms of the bsSMC complex are stuck together. At the end of the coiled coils, the heads are also held together, and with the bridging protein define a single ring-like 'lower compartment' at the base of the complex (Fig. 3a). Then, when ATP binds, there can be a conformational change in the complex which bows the two coiled-coil arms apart, opening a second "upper compartment". The protein "wall" between the two compartments is held closed by the ATPs bound between the two "head" subunits (Fig. 3b); in effect the ATPs act as 'molecular glue" which holds the two heads together while facilitating the opening of the upper compartment.

When the ATPs are hydrolysed the wall between the two compartments of the ATP-bound state is opened, leading to their merging into one large compartment (Fig. 3c). Then, when the ADPs are released (the product of ATP hydrolysis), the two coiled-coils bind together, driving the top portion of the compartment closed, returning the complex to the apo state (Fig. 3a).

In Fig. 3, the (a) to (b) transition involves two distinct steps: nucleotide binding, followed by protein conformational change. We also note that under conditions where ATP and ADP are present at appreciable concentration, there will be competition for binding. Similarly, the transition from (b) to (c) also can be resolved into two successive steps of ATP hydrolysis followed by conformational change. In the next section we will take account of this finer picture of the dynamics of the complex; we will also consider all reverse steps, allowing a complete thermodynamic accounting of energy use by the reaction cycle.

B. Reaction cycle for translocation of SMC along DNA

Given this ATP/protein conformation cycle, we now couple DNA binding to it. There is a highly conserved DNA binding site at the bottom of the upper compartment (this site has been well-characterized in yeast Rad50 [25, 26] and in E. coli MukB). In addition, there is a DNA-binding site at the top of the upper compartment (near the hinge domain) conserved across different SMC family members [24, 27, 28].

In eukaryotes there are, in addition, DNA binding sites in what would correspond to the

bsSMC lower compartment (in the 'wall' region) [29, 30]. A number of SMC complexes have been observed to be able to bind to DNA so as to trap small DNA loops, in the range from 70 nm to 200 nm in size (200 to 600 bp) [31–34], perhaps with DNA bound at two of these binding sites. In the bsSMC case we suppose that there is a mechanism for firmly holding DNA in the lower compartment: this might be an enthalpic interaction, or due to the small size of the lower compartment the interaction could be steric for the case of bsSMC.

Given this, our basic model coupling DNA binding, ATP binding/hydrolysis, and conformational change of the SMC/DNA complex, in terms of a reaction cycle, is sketched in Fig. 4a. We show the same states using molecular models for bsSMC in Fig. 4b (based on Ref. [24]). The upper diagrams are intended to make molecular connectivity and topology clear; gray circles indicate noncovalent DNA-protein interactions. The lower diagrams illustrate practical realizations of molecular conformational changes based on available experimental information [24].

The reaction cycle begins and ends with the enzyme in its apo state "0" (Fig. 4a, leftmost and rightmost states). We assume that in the apo state, a DNA segment is encircled by the lower compartment. Evidence has accumulated that SMC complex function involves entrapment of DNA by protein rings [35–38] so it is plausible that DNA will be in the lower compartment in the apo state, either sterically held, or possibly bound non-covalently to a DNA binding site [39]. Note that in the apo state, the closed coiled-coils prevent binding of DNA to the hinge binding site. The configuration of state 0 can be imagined to be the result of the type of 'loading' reaction associated with initiation of SMC complex activity.

In the apo state 0, there is continuously binding and unbinding of ATP to the two SMC binding sites, and in the conformational state 0, when two ATPs are bound, there is a possibility of a transition to the SMC conformation where the two Walker ATPases are stably bound together (Fig. 4a, state 1). The concomittant opening of the two coiled-coils exposes DNA binding sites in the inside of the upper protein "compartment", permitting capture of a second DNA site as a result of thermal fluctuation of the enzyme complex and the double helix (Fig. 4, transition to state 2). We note that capture of a second DNA site *in cis* is far more likely than *in trans* because of the large local concentration of the former.

Capture of the DNA by the SMC protein upper compartment is key to the mechanochemical cycle. For directional translocation to occur, this second site must be captured preferentially to one side or the other of the initially bound DNA site (i.e., in the "upstream"

direction travelling opposite to the arrowhead shown in Fig. 4a relative to the initially bound DNA site). This may be facilitated by the chirality of the SMC complex, its conformational transitions, and the direction that the DNA is initially "threaded" through and bound the lower compartment: reversing the direction of "threading" will result in loop capture in the opposite direction along the DNA.

Three possible configurations for state 2 of Fig. 4a illustrate how the SMC complex conformational changes may facilitate loop capture to one side of the initially bound DNA site: the "folding" of the lower compartment (state 2, upper right configuration) can align the DNA molecule so that it is "pointed" towards the SMC upper compartment, making loop-capture of upstream DNA more likely than downstream DNA. It should be noted that this folding could occur in either direction (e.g., "into" instead of "out of" the page as sketched in Fig. 4a, state 2) and still direct binding of upstream DNA to the SMC upper compartment.

A variant of this (Fig. 4a, state 2, lower configuration) shows a larger DNA loop of "teardrop" shape with a DNA crossing captured by the complex, a conformation again facilitated by the "folding" of the lower SMC compartment. The interaction shown at the top of the SMC protein compartment in Fig. 4 could be anywhere in the compartment, and might even be simply a steric or quite weak or transient interaction of the DNA loop "trapped" inside the upper compartment.

After capture of the second upstream DNA site, ATP hydrolysis can occur, releasing the Walker ATPases from one another, opening the "gate" between the two compartments, and eliminating the DNA binding sites in the lower compartment. The DNA loop is then free to relax, leading to state 3 of Fig. 4, where the DNA is now held only in the upper compartment, at the upstream position established by the preceding DNA loop capture step.

Finally, ADP release and conformational change back to the apo state push the DNA out of the upper compartment, leading to it being rebound in the lower compartment. Due to the DNA loop capture, the protein complex has translocated upstream on the DNA, or equivalently the DNA has been "pumped" through the SMC protein, by a distance comparable to the contour length of the captured DNA loop.

C. Detailed kinetic-thermodynamic model

To have directional translocation along the DNA, the SMC complex must capture a second site *in cis* preferentially to one side of the initially bound site. The DNA-SMC complex in state 1 can facilitate this by being in a conformation favoring the transition to state 2 (Fig. 4). One must also account for binding, hydrolysis and release of two nucleotides per reaction cycle, along with consideration of ADP binding in the apo state, and the possibility of reversed reactions (ATP synthesis), while maintaining consistency with thermodynamics (the kinetics must be consistent with the free energy released by ATP hydrolysis). We now describe a more complete model (Fig. 5a) and its mathematical formulation, and we show how it can be reduced to a four-state model of the form of Fig. 4.

In Fig. 5a, each state is labeled by the conformation state (0, 1, 2 or 3 as in Fig. 4) followed by nucleotide state (e.g., 2ATP means two ATPs bound; ATP,ADP means one ATP and one ADP bound, and so on). The dashed lines group related states together for reasons which will become apparent shortly. Finally note that Fig. 5 has less topological detail than Fig. 4; SMC-DNA states (0, 1, 2 and 3) of Fig. 5 are meant to correspond to those of Fig. 4.

1. 0↔1: ATP binding and SMC head engagement.

We begin with the apo state, which is able to bind and release ATP, and ADP if it is present. We consider all 16 possible nucleotide combinations (no nucleotide bound, ATP bound, ADP bound, or ADP plus inorganic phosphate bound at each of the two binding sites). Because nucleotide binding/unbinding will occur rapidly in the apo state (the binding sites are exposed) our strategy will be to lump all those states together, into the "0" conformational state (Fig. 5a, left red dashed box). The nucleotide occupation state will be treated as being pre-equilibrated, which makes sense given the rapidity of nucleotide exchange relative to the slower conformational and hydrolysis reactions.

We presume that only in the state where two ATPs are bound (Fig. 5a, 0,2ATP state) can the SMC complex undergo the transition from the rod-like state 0 to the open-coil state 1, where the two ATP-charged heads are engaged ("stuck") together. The net rate for head engagement will depend on the probability of two ATPs being bound, which will in turn

depend on the concentration of ATP (and ADP if sufficiently large). Of course there is also a "disengagement" reverse rate, accounting for the possibility that thermal fluctuations break the two ATP-charged heads apart.

Mathematical formulation - Starting from their apo states, the ATP-binding sites will bind and release ATP at rates typical of ATP-binding enzymes, with an on-rate $k_{\mathrm{on,ATP}} = \gamma_{\mathrm{ATP}}[\mathrm{ATP}]$ where γ is the reaction rate per concentration (units of $\mathrm{M}^{-1}\mathrm{s}^{-1}$) and where [ATP] is in M (mol/litre). For diffusion-limited binding reactions, γ_{ATP} has an upper limit of approximately 10^9 M^{-1} s^{-1} [40], and in practice ATP-binding enzymes typically have ATP binding rates in the range of $\gamma \approx 10^6$ $\mathrm{M}^{-1}\mathrm{s}^{-1}$ [41]. The ATPs can unbind at a rate $k_{\mathrm{off,ATP}}$, which one can expect to be in the range of 10^3 s^{-1} for ATPases with ATP binding site affinities ($K_{\mathrm{d,0,ATP}}$) in the mM range, such as topo II [41]. We take the dissociation constant for an ATP binding site to be $K_{\mathrm{d,0,ATP}} = k_{\mathrm{off,ATP}}/\gamma_{\mathrm{ATP}} = 0.2$ mM. Available data indicates ATP binding affinity $K_{\mathrm{d,0,ATP}} \approx$mM for bsSMC [42].

If ADP is present, it can also bind and unbind at rates $k_{\mathrm{on,ADP}} = \gamma_{\mathrm{ADP}}[\mathrm{ADP}]$ and $k_{\mathrm{off,ADP}}$; similar rates pertain for inorganic phosphate (Pi). In terms of these rates, the dissociation constant for ADP is $K_{\mathrm{d,0,ADP}} = k_{\mathrm{off,ADP}}/\gamma_{\mathrm{ADP}}$. This, and the corresponding affinity for phosphate, are likely less than those for ATP. The dashed box on the left of Fig. 5 is meant to include 16 states corresponding to each Walker A site being apo, or ATP-, ADP- or ADP+Pi-bound. The total probability of all these states is P_0. We set $K_{\mathrm{d,0,ADP}} = K_{\mathrm{d,0,Pi}} = 20$ mM.

Nucleotide on-off rates near physiological concentrations (typically near 1 mM) will be rapid compared to the rates expected for condensin conformational change (Fig. 5 of [24] suggests that the conformational change is rate-limiting), allowing us to treat nucleotide binding in pre-equilibrium. Assuming independent binding of the two nucleotides, if it is given that one is in SMC state 0, the probability of the complex having two ATPs bound is

$$p_{0,2\mathrm{ATP}} = \left(\frac{[\mathrm{ATP}]/K_{\mathrm{d,0,ATP}}}{Z}\right)^2 \tag{1}$$

where $Z = 1 + [\mathrm{ATP}]/K_{\mathrm{d,0,ATP}} + [\mathrm{ADP}]/K_{\mathrm{d,0,ADP}} + [\mathrm{ADP}][\mathrm{Pi}]/(K_{\mathrm{d,0,ADP}}K_{\mathrm{d,0,Pi}})$. The probability of having two ADPs and Pi's bound in state 0 has a similar form,

$$p_{0,2\mathrm{ADP},2\mathrm{Pi}} = \left(\frac{[\mathrm{ADP}][\mathrm{Pi}]/(K_{\mathrm{d,0,ADP}}K_{\mathrm{d,0,Pi}})}{Z}\right)^2 \tag{2}$$

These two probabilities are involved in the two transitions leading out of state 0 in Fig. 5a.

We suppose that the engagement rate of the two ATP-charged heads is k_{eng}. Here we take $k_{\text{eng}} = 2\ \text{s}^{-1}$, *i.e.* a conformational-change time on the order of 500 msec. We presume that the reverse head-opening rate k_{dis} is slower: $k_{\text{dis}} = k_{\text{eng}} e^{-\beta \varepsilon_{\text{eng}}}$. Here $\varepsilon_{\text{eng}} = 4k_B T$ indicates the head engagement ("sticking") free energy, since the ratio of the forward to reverse engagement rates is related to the free energy difference between engaged and disengaged states through a Boltzmann factor. This type of free energy accounting of reverse relative to forward rates will recur as we construct the reaction cycle and is essential to understanding its thermodynamics.

2. 1↔2: reversible DNA loop capture.

Once the upper compartment is open, the DNA binding site at the top of it is exposed, allowing a DNA loop to be captured (Fig. 5a, state 2). Due to the left-right asymmetry of the protein complex (all SMC complexes possess this asymmetry), we can expect a loop to form "from the left" via capture of upstream DNA or "from the right" with unequal rates. Note that the DNA "track" is unpolarized, so that left-right symmetry-breaking of the SMC complex is essential to ensure processive translocation. We suppose that capture of an upstream DNA loop dominates, and we ignore "looping from the right" (this could be included but will not essentially change the model). DNA loop formation of the transition from state 1 to state 2 is reversible, since state 2 is held together by non-covalent bonds and can be destabilized by tension in the DNA.

The forward rate k_1 requires a loop of DNA to be formed, and as a result depends on the mechanics of DNA (its persistence length $A = 50$ nm, or perhaps the slightly shorter effective persistence length of nucleoid-associated-protein-coated DNA [43, 44] as found in a bacterial cell) and also on any tension f_{DNA} that might be present in the DNA itself. If DNA tension is present, it will compete with bending energy to select a smaller loop size [46], as well as slowing down the forward loop-capture rate. Two key points emerge from this transition: the step size for translocation will be on the order of the size of the DNA loop that is trapped by the SMC (the position of the second, upstream interaction site between DNA and SMC); and, increased DNA tension will reduce the step size. Again, there is the possibility of a reverse transition whereby the DNA loop is released: this reverse transition is likely to be increased in rate by increasing DNA tension.

Mathematical formulation - The most likely size of loop ℓ that will form at low values of $f_{\rm DNA} < 0.1$ piconewtons (pN) is comparable to the $\ell \approx 200$ nm which is the peak of the DNA cyclization J-factor distribution [45]. As mentioned At larger forces, competition between force and bending will select a loop size of $\ell = \sqrt{Dk_BTA/f_{\rm DNA}}$, where D is a numerical constant equal to about 14; the energy of this loop is $\varepsilon_{\rm loop} = 2\sqrt{k_BTDAf}$ [46], which takes into account the mechanical work done against DNA tension $f_{\rm DNA}$. A suitable (free) energy and associated loop size describing the loop including the low-force limit is

$$\begin{aligned} \varepsilon_{\rm loop} &= 2\sqrt{Dk_BTA(f_0 + f_{\rm DNA})} \\ \ell &= \sqrt{Dk_BTA/(f_0 + f_{\rm DNA})} \end{aligned} \tag{3}$$

where $f_0 = 0.1$ pN provides a sensible low-force limit roughly corresponding to zero-force cyclization. As force is increased, the optimal loop size decreases, and loop energy increases. We take a fraction ϕ ($0 < \phi < 1$) of the energy $\varepsilon_{\rm loop}$ as the energy barrier for the forward reaction (1→2), indicating $k_1 = k\exp[-\beta\phi\varepsilon_{\rm loop}]$. Here k is a thermal fluctuation rate for an object of $d = 10$ nm length, or $k = k_BT/(\eta d^3)$; we take $\phi = 1/2$.

We need to take into account the fact that if the most probable-sized loop considered above is too small (i.e., if the force is large), it will not be able to be captured by the SMC. To include this effect we add a DNA loop-capture probability $p = 1/(1 + e^{(\ell_0-\ell)/\delta})$ with $\delta = 30$ nm and $\ell_0 = 80$ nm, to shut off the loop-capture rate when ℓ drops below the lower limit of loop size compatible with capture by the SMC, by more than δ:

$$k_1 = kpe^{-\beta\phi\varepsilon_{\rm loop}} \tag{4}$$

Although simple, over a range of forces up to a few pN this model describes the barrier loop energy cost increase and loop size decrease as $f_{\rm DNA}$ is increased, including the inability of the SMC to capture loops which are smaller than ℓ_0. The numerical parameter D is determined by the loop bending geometry, so the precise details of the DNA conformation captured by an SMC complex can be modeled by varying the value of D, ℓ_0 and δ and ϕ.

Reversal of loop capture can occur as a result of thermal breaking of the upper-compartment DNA-SMC interaction, *i.e.*, $k_1' \propto e^{-\beta\varepsilon_{\rm bind}}$. This takes into account the binding energy of the SMC to the DNA loop, but we must also include the remainder of the loop-formation free energy which now plays the role of a tension-dependence driving the DNA loop off the complex:

$$k_1' = kpe^{-\beta\left[\varepsilon_{\rm bind}-(1-\phi)\varepsilon_{\rm loop}\right]} \tag{5}$$

where $\varepsilon_{\text{bind}}$ is the energy of DNA-loop-binding; we take $\varepsilon_{\text{bind}} = 15k_BT$. The bending energy ensures that the ratio of the forward and reverse rates gives the Boltzmann factor describing the relative probabilities of states 1 and 2 (with ATP bound), corresponding to a free energy difference $F_{2,\text{ATP}} - F_1 = \varepsilon_{\text{loop}} - \varepsilon_{\text{bind}}$.

3. 2↔3: ATP hydrolysis, phosphate release, and loop release.

When the ATPs are hydrolysed, the two heads are released from one another, and the wall between the upper and lower compartments is opened. This requires two successive hydrolysis steps with two intermediates with one and two ADPs bound (2,ATP,ADP and 2,2ADP of Fig. 5a, respectively, inside the green dashed box). We take the forward rate $k_{\text{hyd}} = 20\ \text{sec}^{-1}$ [41]. Despite being exceedingly slow under almost any conceivable experimental condition, we will consider reverse ATP synthesis steps.

ATP hydrolysis causes the two Walker ATPase heads to be released from one another, which we presume eliminates the assumed DNA binding capacity of the lower compartment. The DNA site bound in step 0 is released, allowing the DNA to pop through the now-opened "gate" between the lower and upper compartments (Fig. 5a, state 3). This is the translocation step, since it moves DNA contour (arrow) that was on the left side of the SMC complex in state 0, to the right side of the SMC.

The result is the 2-ADP-bound state 3 which has DNA held only in the upper compartment (Fig. 5a, state 3,2ADP); this state is likely to be highly transient since the ADPs have relatively low affinity and are not being held in place by the engaged heads. In steady state it will turn out that we can eliminate the two ATP hydrolysis steps to obtain a single pair of rates linking conformational state 2 to state 3.

We note that the DNA loop capture and release process involves formation of DNA crossings, which can generate supercoiling of the DNA template. However, our model does not require crossings of a specific sign, and by forming crossings of opposite signs, supercoiling of the DNA can be avoided. On the other hand, for SMC complexes which trap DNA loops of specific or preferred chirality, it is conceivable that supercoiling of the DNA template could be generated along with translocation.

Mathematical formulation - After setting the ATP hydrolysis rate, thermodynamics di-

cates the synthesis rate $k_{\rm syn}$ via

$$\frac{k_{\rm syn}}{k_{\rm hyd}} = \frac{K_{\rm d,2,ADP} K_{\rm d,2,Pi}}{K_{\rm d,2,ATP}} \frac{[{\rm ATP}]_{\rm eq}}{[{\rm ADP}]_{\rm eq}[{\rm Pi}]_{\rm eq}} \tag{6}$$

where the $K_{\rm d,2,X}$ refer to the binding affinities of species X to the SMC in protein state 2. These affinities are in general different from the $K_{\rm d,0,X}$ (the protein is in a different conformation) but they are likely similar in order of magnitude (the Walker A nucleotide binding site is in similar conformation).

Eq. (6) has a simple interpretation: the free energy released by ATP hydrolysis in solution (given by the log of the ratio of equilibrium concentrations) is slightly shifted by binding free energy (the log of the ratio of the affinities). Under almost any conceivable enzyme conditions we will have $k_{\rm syn} << k_{\rm hyd}$ since the solution equilibrium of ATP is so strongly skewed towards hydrolysis, as will be made quantitatively clear below. A kinetic derivation of Eq. (6) is given in Supplementary Data. We take $K_{\rm d,2,X} = K_{\rm d,0,X}$ for $\rm X = ATP, ADP, Pi$ due to the lack of detailed information and also since there is relatively little effect of their precise values on the results of our model.

In the above, [Pi] and $[{\rm Pi}]_{\rm eq}$ are concentrations and equilibrium concentrations of phosphate); phosphate is involved in ATP synthesis/hydrolysis, and affects ATP/ADP free energy balance. Note that we suppose that phosphate release/binding is coincident with ADP release and binding. If phosphate release/binding is indicated by future experimental data to be at a different point in the cycle of Fig. 5a than ADP release, it is straightforward to relocate, but this will have no effect under *in-vivo*-like reactions conditions where the rate of ATP synthesis (which involves phosphate binding and therefore [Pi]) are tiny relative to that of hydrolysis.

Once hydrolysis has occurred, we presume a forward "opening" rate $k_{\rm open} = 200$ sec^{-1}. Given that this involves opening and relaxation of the DNA loop and release of DNA binding, we write $k_{\rm close} = k_{\rm open} e^{\beta[\varepsilon_{\rm bind} - \varepsilon_{\rm loop} - \varepsilon_{\rm open}]}$ where $\varepsilon_{\rm open} = 10 k_B T$ is the additional free energy driving the SMC-DNA-ADP complex towards the open state.

4. 3↔0: ADP release.

The final step of the reaction cycle is the closing of the upper compartment (reformation of the rod-like conformation of the SMC coiled-coils) which resets the SMC back to confor-

mational state 0 and pushes the DNA back into the lower compartment where it rebinds. Relative to the initial state, the SMC complex is now translocated along DNA by an amount of approximately ℓ, the size of the loop captured during the 1→2 transition. The rate of this process is given by $k_3 = 10$ sec^{-1}. We include a rate k_3' for the reverse of this process, which we will set shortly using energy dissipation considerations.

D. Reduction to the four-state model

We can simplify the model of Fig. 5a to a four-state cycle corresponding to the four major conformational states described earlier (Fig. 4). We take advantage of the rapid equilibration of nucleotide-binding states in SMC state 0 (Fig. 5a, left red dashed box) and consider them together, with net probability P_0. Given pre-equilibration of nucleotide binding, the transition from the block of 0 states to the 1 state occurs at rate $k_0 = p_{0,2\mathrm{ATP}} k_{\mathrm{eng}}$. Note that k_0 connects ATP concentration to forward cycling: k_0 increases linearly as [ATP] is increased from zero, and saturates at high [ATP]. The reverse transition is $k_0' = k_{\mathrm{dis}} = k_{\mathrm{eng}} e^{-\beta \epsilon_{\mathrm{eng}}}$.

In steady state, a second simplification of the scheme of Fig. 5a follows from combining the transitions linking 2,2ATP to 3,2ADP into a single transition, by "integrating out" the 2,ATP,ADP and 2,2ADP states (Fig. 5a, right green dashed box). This can be done by writing the steady-state balance equations for the four states, and then eliminating the probability of the 2,ATP,ADP and 2,2ADP states. The resulting rates are

$$\begin{aligned} k_2 &= 2(k_{\mathrm{hyd}})^2 k_{\mathrm{open}}/\Delta \\ k_2' &= 2(k_{\mathrm{syn}})^2 k_{\mathrm{close}}/\Delta \\ \text{where} \quad & \Delta = (2k_{\mathrm{syn}} + k_{\mathrm{open}})(k_{\mathrm{syn}} + k_{\mathrm{hydr}}) - k_{\mathrm{syn}} k_{\mathrm{hydr}} \end{aligned} \tag{7}$$

One should keep in mind that there is still probability associated with the now "hidden" 2,ATP,ADP and 2,2ADP states

$$\begin{aligned} P_{2,\mathrm{ATP,ADP}} &= [2k_{\mathrm{hydr}}(2k_{\mathrm{syn}} + k_{\mathrm{open}})/\Delta] P_{2,2\mathrm{ATP}} \\ &+ [2k_{\mathrm{syn}} k_{\mathrm{close}}/\Delta] P_{3,2\mathrm{ADP}} \\ P_{2,2\mathrm{ADP}} &= [4(k_{\mathrm{hydr}})^2/\Delta] P_{2,2\mathrm{ATP}} \\ &+ [k_{\mathrm{close}}(k_{\mathrm{hydr}} + k_{\mathrm{syn}})/\Delta] P_{3,2\mathrm{ADP}} \end{aligned} \tag{8}$$

and that $P_2 = P_{2,2\mathrm{ATP}} + P_{2,\mathrm{ATP,ADP}} + P_{2,2\mathrm{ADP}}$. Derivations of these formulae are given in Supplementary Data.

Once these steps are done, the model is expressed as shown in Fig. 5b, in terms of the four SMC-DNA conformational states of Fig. 4. This reduced model is exactly equivalent to the detailed model of Fig. 5a (as long as we consider only steady-state cycling); it allows us to focus on the experimentally relevant DNA-tension-dependence associated with loop capture in k_1 and k_1' and the ATP concentration dependence in k_0.

E. Energy consumption during the reaction cycle

In steady state, the net flux (total forward minus backward transitions per unit time) between each pair of adjacent states is equal, and we have the thermodynamic relation

$$\frac{k_0 k_1 k_2 k_3}{k_0' k_1' k_2' k_3'} = \left(\frac{[\mathrm{ATP}]/[\mathrm{ATP}]_{\mathrm{eq}}}{([\mathrm{ADP}]/[\mathrm{ADP}]_{\mathrm{eq}})([\mathrm{Pi}]/[\mathrm{Pi}]_{\mathrm{eq}})} \right)^2 \tag{9}$$

This relation must hold since the energy dissipated during one cycle must correspond to the free energy released by hydrolysis of 2 ATPs. Eq. (9) indicates that the cycle will run "forward" (towards ATP hydrolysis) when ATP concentration is sufficiently high relative to ADP concentration, that the cycle will run backwards under conditions of sufficient excess ADP/phosphate, and that the cycle will cease when equilibrium is reached (concentrations equal to their equilibrium values). "Integrating out" the 2,ATP,ADP and 2,2ADP states does not alter the detailed balance relation (9) since $k_2/k_2' = k_{\mathrm{hyd}}^2 k_{\mathrm{open}}/(k_{\mathrm{syn}}^2 k_{\mathrm{close}})$.

The thermodynamic constraint (9) sets the as-yet-undetermined rate k_3':

$$\begin{aligned} k_3' \;=\; & k_3 p_{0,2\mathrm{ADP},2\mathrm{Pi}} e^{\beta[\varepsilon_{\mathrm{eng}}+\varepsilon_{\mathrm{open}}]} \\ & \left(\frac{K_{\mathrm{d},2,\mathrm{ATP}}}{K_{\mathrm{d},0,\mathrm{ATP}}} \frac{K_{\mathrm{d},0,\mathrm{ADP}}}{K_{\mathrm{d},2,\mathrm{ADP}}} \frac{K_{\mathrm{d},0,\mathrm{Pi}}}{K_{\mathrm{d},2,\mathrm{Pi}}} \right)^2 \end{aligned} \tag{10}$$

This form of this reverse rate makes physical sense: the leading term selects only the subpopulation of state 0 with ADP and Pi bound for the transition in the direction of k_3', and the Boltzmann factor provides precisely the energy cost of opening the closed "hairpin" without the ATP-closed gate. The final terms take into count the free energy difference between cofactors bound to state 0 and state 2. An alternative computation of (9) is included in the Supplementary Data.

F. Free energy release from ATP hydrolysis and equilibrium nucleotide concentrations

One can consider as input to the model the concentrations of phosphate and nucleotides, i.e., [ATP], [ADP] and [Pi] as prepared in a given experimental situation or as occur *in vivo*. Our model also requires the equilibrium concentrations of these reactants given those inputs, which is determined by the dissociation constant K for ADP/ATP equilibrium, via $[\mathrm{ADP}]_{\mathrm{eq}}[\mathrm{Pi}]_{\mathrm{eq}} = K[\mathrm{ATP}]_{\mathrm{eq}}$. Experimental data [47] indicate $K = e^{16}$ M (see Supplementary Data); given that in vitro or in vivo [ATP] is far above this allows us to use the first term of an expansion in $[\mathrm{ATP}]/K$ of the exact result for $[\mathrm{ATP}]_{\mathrm{eq}}$ (see Supplementary Data)

$$[\mathrm{ATP}]_{\mathrm{eq}} = \frac{([\mathrm{ADP}] + [\mathrm{ATP}])\,([\mathrm{Pi}] + [\mathrm{ATP}])}{K} + \mathcal{O}\left(\frac{[\mathrm{ANP}]^3}{K^2}\right) \tag{11}$$

with $[\mathrm{ADP}]_{\mathrm{eq}} = [\mathrm{ADP}] + [\mathrm{ATP}] - [\mathrm{ATP}]_{\mathrm{eq}}$. and $[\mathrm{Pi}]_{\mathrm{eq}} = [\mathrm{Pi}] + [\mathrm{ATP}] - [\mathrm{ATP}]_{\mathrm{eq}}$.

III. RESULTS

A. Steady-state cycling rate

We can compute the steady-state probabilities P_i of each of the states of Fig. 5b, and the steady-state cycling and DNA translocation rates. In steady state there must be equal flux into and out of each state, giving the set of equations

$$(k'_{i-1} + k_i)P_i = k_{i-1}P_{i-1} + k'_i P_{i+1} \tag{12}$$

where the subscripts are considered modulo 4, and where state 2 represents the 2,2ATP state (as in Fig. 5b). In addition we have the constraint that the total probability of all states sums to unity, or $P_0 + P_1 + P_{2,\mathrm{2ATP}} + P_{2,\mathrm{ATP,ADP}} + P_{2,\mathrm{2ADP}} + P_3 = 1$ (recall that $P_{2,\mathrm{ATP,ADP}}$ and $P_{2,\mathrm{2ADP}}$ state probabilities are determined by Eq. 8).

This system of equations is analytically solvable, although rather lengthy and uninformative formulae are obtained. It can be shown that the flux (cycling rate k_{kcycle}) through the cycle (any of $k_i P_i - k'_i P_{i+1} = k_{\mathrm{cycle}}$ for $i = 0, \cdots, 3$) is proportional to $([\mathrm{ATP}][\mathrm{ADP}]_{\mathrm{eq}}[\mathrm{Pi}]_{\mathrm{eq}})^2/([\mathrm{ATP}]_{\mathrm{eq}}[\mathrm{ADP}][\mathrm{Pi}])^2 - 1$, i.e., that nucleotides and phosphates must be perturbed from their equilibrium values for a cycle (nonequilibrium) to occur (see Supplementary Data). In any case the cycling rate is straightforward to compute from the exact equations (12).

B. Translocation rate

The cycling rate for the SMC-DNA complex k_{cycle} is not yet the translocation rate. The forward velocity of the SMC along its DNA template is given by ℓk_{cycle}, but if there is an load - a force impeding the SMC's progress along DNA of f_{load} (Fig. 6a) - one can expect some slippage of the SMC backwards along the DNA. Notably, for simple translocation, the load f_{load} acts independently of the tension in the DNA template f_{DNA}, so the slippage can be treated as a process which is essentially decoupled from the loop-capture-translocation. The simplest model is one where there is slippage of the DNA from one of its binding sites at some point during the translocation cycle.

We suppose that such events are associated with a slipping length Δ and that the barrier opposing these slip events is on the order of $\varepsilon_{\text{slip}}$. A simple force-driven barrier-crossing model gives a reverse slipping velocity of

$$v_{\text{slip}} = k\Delta \left[e^{-\beta(\varepsilon_{\text{slip}} - \Delta f_{\text{load}})} - e^{-\beta(\varepsilon_{\text{slip}} + \Delta f_{\text{load}})} \right] \tag{13}$$

where the two terms describe random slip in the load force direction and against it; at $f_{\text{load}} = 0$, slipping in either direction is equally (im)probable, giving $v_{\text{slip}} = 0$. The constant Δ is the range of deformation of the DNA-SMC bound state that is required to break their chemical contact, likely of nm dimensions; here we take $\Delta = 2$ nm. We take the slipping energy $\varepsilon_{\text{slip}} = 14 k_B T$. Given this slip velocity, the net translocation velocity is

$$v = \ell k_{\text{cycle}} - v_{\text{slip}} \tag{14}$$

Since translocation velocity is observable in single-molecule SMC-DNA experiments, we now examine its behavior for reasonable choices of parameters. The solid line in Fig. 6b shows the (exact) translocation velocity as a function of ATP concentration at zero load force and DNA tension of $f_{\text{DNA}} = 0.2$ pN, showing how the translocation velocity saturates at high ATP concentration. For the solid line, ADP concentration is held fixed at 1 μM; the dashed line shows how translocation is slowed when ADP is increased to 0.1 mM, due to the decreased free energy drive associated with the increased ADP concentration (see Eq. 9).

Fig. 6c shows the (exact) translocation velocity as a function of template DNA tension f_{DNA} and for fixed ATP concentration of 1 mM, fixed ADP concentration of 1 μM, and zero load force, showing how translocation is slowed down by tension in the DNA being

translocated along, and then is finally essentially stopped, as discussed above. The translocation crucially depends on the rate k_1, which is a product of the loop-formation Boltzmann factor and the loop capture probability p (Eq. 4). Fig. 6d shows how these two quantities vary with DNA tension for the same ATP and ADP concentrations; the probability for loop formation dramatically decreases (Fig. 6d, solid curve), shutting off translocation. There is also a weaker suppression effect (Fig. 6d, gray dashed curve) generated by the decrease in loop size (Fig. 6c, gray dashed curve).

Fig. 6e shows the (exact) evolution of the probabilities of the four states as DNA tension is increased (for the same ATP and ADP concentrations as in Fig. 6c-d): at low tension, state 0 is highly probable due to the conformational change rate being rate limiting. Then, as force is increased, DNA loop capture becomes rate-limiting, resulting in a decrease of P_0 and an increase of P_1. For the parameters used here, the probabilities of P_2 and P_3 are similar; they start at 5 to 10% at low force and then are both pushed to zero when the tension in the DNA becomes large enough to stall the motor, enforcing occupation of state 1 with high probability.

Finally, Fig. 6f shows the translocation velocity as a function of load force, which eventually not only slows the SMC translocation velocity but can reverse it, pushing the enzyme back by inducing a high rate of slipping events. The relatively large stall via load force may be difficult to observe, given that most experimental arrangements to apply a load force will also apply DNA tension, which will quench translocation at a low force.

Key predictions of our model follow from the loop-capture mechanism and are apparent from Fig. 6c. First, stalling of translocation should occur for subpiconewton forces, well below the few-pN stall forces typically observed for cytoskeletal ATP-powered motors such as kinesins. Second, the average step size of SMCs should be reduced by increased DNA tension (the average translocation step size should approximately correspond to the loop size shown in Fig. 6c). Third, the DNA tension cutoff for translocation with DNA tension should be adjustable by changing SMC complex size (more precisely, by adjusting the size of DNA loops that can be captured). It may be possible to adjust the cutoff l_0 (or perhaps the DNA loop-shape dependent constant D) by changing the length of the SMC coiled-coils, with longer coils leading to a cutoff of translocation at lower forces, and with shorter coils leading to a cutoff at higher forces.

In Supplementary Data we include additional data for the extrusion velocity ignoring

slippage, showing dependence on ATP/ADP ratio (Supplementary Fig. 1), total nucleotide concentration (Supplementary Fig. 2), versus force for varied total nucleotide concentration (Supplementary Fig. 3), and for varied phosphate dissociation constants (Supplementary Fig. 4). We also include data for the probabilities of the four states as a function of ATP/ADP ratio (Supplementary Fig. 5) and DNA tension (Supplementary Fig. 6).

C. Approximation of small reverse rates and rapid head opening: Michaelis-Menten-like behavior

In an experimentally-relevant limit, the complete model of Fig. 5 (numerical steady-state results of Fig. 6) (12), has a relatively simple analytical solution which is physically informative. Given the parameters described above and for ATP concentrations in the physiological range (near mM), the reverse rates of Fig. 5b are negligible relative to the forward ones, with one exception. The transition from state 1,ATP back to state 0,ATP is strongly suppressed by the free energy favoring head engagement; ATP synthesis is much slower than ATP hydrolysis when ATP and ADP are anywhere near mM concentrations, the transition from the opened state 3,2ADP back to 2,2ADP is supressed by the free energy cost of forcing the gate closed in the absence of ATP, and the transition from 0,ADP to 3,ADP is similarly supressed by the cost of forcing the "hairpin" open.

The only transition which can have backward steps that can compete with forward steps under physiologically-relevant conditions is the DNA-loop-capture transition from 1,ATP to 2,2ATP (k_1' of Fig. 5). The balance between these forward and reverse transitions can be changed using DNA tension, making this step crucial to controlling the cycling rate. In fact, eliminating all of the reverse steps except k_1' leads to nearly identical results to the model with all reverse steps, and also allows a simple analysis of the model's behavior.

We can also make the approximation that little time is spent in the 2,ADP,ADP and 2,2ADP states relative to the 2,2ATP state, which results from the head opening rate (k_{open}) being fast compared to the ATP hydrolysis rate (k_{hyd}). Given the relatively slow rate of ATP hydrolysis, this is likely, and allows the further approximation that $P_2 \approx P_{2,\text{ATP}}$; in our approximate treatment we take $P_2 = P_{2,\text{ATP}}$.

Given these approximations ($P_2 = P_{2,\text{ATP}}$, all reverse rates set to zero except k_1', see

Supplementary Data) the cycling rate is

$$k_{\text{cycle}} = \frac{k_0}{1 + (B + C/k_1)k_0} \tag{15}$$

where $B = 1/k_2 + 1/k_3$ and $C = 1 + k_1'/k_2$ are used to write the simplified final expression. We note that k_0 and k_1 are easily related to experimental control parameters; k_0 is controlled by ATP concentration, and k_1 is controlled by the tension in the DNA (f_{DNA}).

The cycling rate has a Michaelis-Menten form with respect to k_0: for low k_0 (achieved for sufficiently low ATP concentration), the cycling rate approaches k_0, but for sufficiently large k_0, the cycling rate saturates at $k_{\text{cycle,max}} \to 1/(B + C/k_1)$, reflecting that transitions other than 0→1 are rate-limiting for the reaction cycle when k_0 is sufficiently large (*i.e.*, $k_0 \gg k_{0,m} = 1/[B + C/k_1]$).

For fixed k_0 (ATP concentration) the behavior of k_{cycle} with f_{DNA} (via k_1) is interesting. As the tension is increased above $f_0 \approx 0.1$ pN, k_1 starts to drop, and eventually k_1 will become rate-limiting, with $k_{\text{cycle}} \to k_1/C = k_1/(1 + k_1'/k_2)$. We note that this effect, of tension in the template (f_{DNA}) controlling the rate of motion of a motor, is rather unusual. We also note that f_{DNA} is not, strictly speaking, a load on the motor; instead it is controlling one of the forward reaction steps rather than driving a reverse step. Interestingly, f_{DNA} can slow SMC translocation down (essentially stopping the enzyme for large enough f_{DNA}) but it cannot *reverse* the direction of motion of the complex down DNA in the way a load force applied directly to the enzyme might be expected to. This type of dependence of SMC translocation on f_{DNA} - translocation quenching but without reversal - is perhaps the clearest hallmark of this type of loop-capture model.

IV. DNA LOOP EXTRUSION

Above we have shown that a loop-capture mechanism is thermodynamically capable of generating translocation of SMC complexes along a thermally fluctuating DNA molecule. It has been proposed that SMC complexes are capable of "extruding" loops of DNA, for example by the pushing apart of two motile elements [3, 48], and recent single-molecule experiments have observed such behavior [23]. We now discuss a few ways whereby the model of Sec. II could be involved in active DNA loop extrusion.

A. Two-translocator model

The first possibility is that two translocators of the type described in Sec. II could be physically linked together (Fig. 7a) so as to have a DNA loop pushed out. To have loop extrusion, the two translocators must be loaded in a specific orientation. If the DNA is topologically entrapped inside the protein "ring" of each translocator, the orientations of the two translocators will be preserved following the initial "loading". For this model, the velocity of loop extrusion is the sum of the translocation velocities of the two translocators, or in terms of the translocation velocity of Eq. 14, $v_{\mathrm{loop-extrusion}} = 2v$.

B. Model for loop extrusion by a single translocator

A second possibility is that a DNA loop could be extruded by a single translocator of the type described in Sec. II (Fig. 7b) [24]. This requires only a slight variation on the translocator model; instead of trapping just one DNA in the lower compartment (Fig. 5, state 0 and 1), we suppose that two DNAs (for instance, the base of a DNA loop) can be trapped (Fig. 7b, state 0). Starting from state 0, a small DNA loop is captured as before in the upper compartment (Fig. 7b, state 1), and then ATP hydrolysis merges the two compartments, allowing transfer of the length in the smaller (upper) DNA loop into the larger (lower) DNA loop. The reaction model for this single-translocator loop-extrusion process is just the same as for translocation (Eq. 14), *i.e.*, $v_{\mathrm{loop-extrusion}} = v$.

The model as shown in Fig. 7b moves the upper DNA strand through the lower compartment, but the lower strand stays bound to the protein through the cycle. If iterated, this reaction will extrude a loop, but in an asymmetric manner, with one loop boundary site permanently bound. With two DNAs bound in close proximity in the lower compartment, it is likely that they will exchange binding sites periodically (likely rapidly given the high effective concentration of DNA within the lower compartment [49]), which will result in symmetric loop extrusion, still with the loop extrusion velocity equal to the DNA translocation velocity, or $v_{\mathrm{loop-extrusion}} = v$.

C. Asymmetric single-translocator model with permanently bound DNA

A variation of the single-translocator model might be one where a flanking DNA is bound to the *exterior* of the SMC protein complex through the entire reaction cycle (Fig. 7c), for example via the DNA binding sites located at the walls of the lower compartment in condensin and SMC5/6 [29, 30]. If so, then translocation to the left would result in asymmetric loop extrusion, as has been observed for yeast condensin [23]. While this model provides a simple explanation for how SMC translocation can result in DNA looping, asymmetric loop extrusion is likely to perform poorly in chromosome compaction as it will tend to leave stretches of uncompacted DNA [18].

V. DISCUSSION

We have presented a quantitative statistical-mechanical kinetic model for SMC complexes based on structural and catalytic information for the *B. subtilis* bsSMC complex (Fig. 2). Our model follows the DNA loop-capture/protein conformational change scheme proposed in Ref. [24] (Fig. 3-5), and our theoretical analysis now shows how it can generate translocation (and loop extrusion) in reasonable accord with available experimental information. Although our model contains a number of parameters, we have chosen rather natural values for them, with the result that essentially the translocation involves DNA-tension-dependent steps of from 50 nm to 150 nm in size, with a choking off of translocation for DNA tensions of $f_{\mathrm{DNA}} \approx 1$ pN (Fig. 6b).

A. DNA substrate-tension f_{DNA} dependence of SMC translocation velocity

A key feature of our model is its dependence of translocation velocity on tension in the DNA 'track' that the SMC is moving along: this is unique for a molecular motor (note that myosins are not sensitive to the tension in their F-actin tracks): this reflects the essential dependence of the mechanism of our model on DNA looping, which is of course quenched by DNA tensions exceeding the pN scale [46]. Our model is an example of a novel molecular motor system, for which translocation essentially depends on bending and consequent looping of the track itself.

We emphasize that this dependence of translocation on substrate tension provides an

important experimental test of the model: tension in a DNA of more than a few pN should stop SMC translocation. Recently translocation of yeast condensin has been observed along DNA molecules under some tension [22], but in that experiment the applied tension was not known (however it was likely less than 1 pN). Experiments observing compaction of DNA by various types of SMC complexes have observed compaction reactions which are stalled by forces exceeding roughly 1 pN [31–34, 50], also suggestive of a loop-capture mechanism. As this paper was in preparation, quenching of loop extrusion by DNA tension ≈ 1 pN was reported for yeast condensin, providing good evidence for a DNA-loop-capture mechanism underlying DNA translocation by that enzyme complex [23].

If this relatively low (≈ 1 pN) stall force (small relative to the multi-pN stall forces of common motors like kinesins and myosins) does turn out to be a general feature of DNA-translocating SMCs, one might ask why such a "weak" motor system has been selected to play key roles in chromosome dynamics. One reason is likely that ≈ 1 pN forces are high enough to extend and transport DNA (and to bias strand transfer by type-II topoisomerases), but low enough to not significantly disturb transcription factors and DNA packaging factors non-covalently bound to the double helix [51–53]. Thus "weak" SMCs offer the feature that they can reorganize DNA without disturbing patterns of gene expression or ≈ 10 nm-scale DNA packaging machinery. Another way to look at the force scale for SMC stalling is that it is tuned so that SMCs act on DNA organization at relatively large (roughly DNA/chromatin persistence length and larger, or > 10 nm) length scales.

B. Load force f_{load} and SMC translocation velocity

The substrate tension is to be distinguished from a load force applied to the SMC complex, which directly acts against translocation (Fig. 6a inset). Substrate tension is not able to push an SMC complex backwards, hence the quenching of the forward velocity (Fig. 6b). In contrast, a load force can slow the enzyme to zero velocity and even push it backwards, essentially by forcing the enzyme to unbind from its DNA substrate (Fig. 6d).

For the parameters we have chosen, the zero-velocity point is reached at a load force of about 3.5 pN (Fig. 6d), but this will undoubtedly be a rather complicated story as this is dependent on the details of the binding of SMC to the DNA double helix through the enzyme reaction cycle. Still, we expect the basic distinction between response to substrate

tension (no reversal at large forces) and load force (reversal at large forces). shown in Figs. 6b and 6d to remain, whatever the details are of these curves.

We note that in DNA compaction assays against applied force [23, 31–34] (Fig. 8a) the applied force acts as both substrate tension *and* as a load force (the load force most likely being a fraction of the substrate tension) potentially complicating the interpretation of such experiments. Experiments of the sort of Ref. [22] studying translocation as a function of substrate tension, with a separately controlled load force (for example through attachment of a handle directly to the translocating SMC to apply a load directly) could decouple these two control parameters (as sketched in Fig. 6a).

C. Further refinements of the model

Two steps of the proposed reaction scheme may rely on precisely balanced SMC-DNA interactions. First, the formation of DNA loops to be captured by SMC is rate-limiting at high DNA tension in the simplest version of our DNA-segment-capture model (Fig. 6d). SMC complexes may facilitate DNA looping as well, for example by stabilizing early intermediates of the DNA looping reaction via specific SMC-DNA contacts. If so, then passing through the transition state may require only a reduced level of DNA bending rather than full looping, which could significantly reduce the energy barrier for the transition from states 1 to 2.

Conceivably, SMC may locally 'melt' (strand-separate) a segment of the DNA double helix by binding single-stranded DNA and thus favour strong DNA curvatures [54] (bsSMC has long been known to bind ssDNA rather well without a clearly understood function). SMCs may also rely on other cellular proteins (such as the bacterial DNA-bending protein HU or nucleosomes) to generate strongly bent DNA *in vivo*. In our model, the quantities ϕ and D can be used to roughly take these phenomena into account, by reducing the energy barrier to loop formation (our choices of the DNA bending-energy parameter D and the forward/reverse rate-splitting parameter ϕ consider a completely unconstrained loop, with the effect of DNA bending free energy split equally between forward and reverse rates). For small ϕ or D values, the energy cost for the formation of a DNA loop is reduced, speeding up SMC translocation especially for larger DNA tensions (> 1 pN).

A second critical step of the model is the return of the ADP-bound state (Fig. 5, state 3)

to the apo state (0) by the closure of the upper compartment and the transfer of the DNA double helix from the top of the upper compartment all the way into the lower compartment. Zipping up of the SMC arms must preferentially initiate at the top end and happen quickly so that loss of translocated DNA by free sliding through the closing upper compartment is limited. Alternatively, the DNA transfer may occur in discrete steps with DNA being fixed at each step via specific SMC-DNA contacts. We envision that this transition is critically dependent on properly folded SMC arms and blocked by SMC arms with aberrant lengths or local flexibility [42].

A less critical aspect of the model is the precise order in which, following ATP hydrolysis, ADP is released, in relation to the closing of the upper compartment (Fig. 5a, transition from state 3,2ADP to 0,2ADP). It is straightforward to reverse this order, i.e., to have ADP release precede the upper compartment closing; this will not change any of the qualitative results of the model, but will only require addition of an additional kinetic step. Similarly, the two ATP hydrolysis steps could proceed at different rates; again, this can be incorporated without any qualitative changes to the model.

D. Application of the model to different types of SMC complexes

We have focused on bsSMC in this paper since we have more information about structural, enzymatic and *in vivo* function than any other SMC complex. We anticipate that our model will be applicable to other SMC complexes, be they eukaryote, bacterial, or archaeal. We now discuss some details of how our model might apply to specific types of SMC complexes.

1. Eukaryotic condensin.

Eukaryotic condensin, at present, is thought to be loaded essentially randomly onto DNA (without specific binding sites), and then to subsequently act to compact DNA. The most simple way for this to occur is by binding, and then loop extrusion (Fig. 8b; note that the enzyme shown might be one or two SMCs as discussed in Sec. IV). If multiple extruders bind, they can 'stack up' to produce more robust compacted loops [3, 18, 19].

A recent experiment has observed translocation of yeast condensin along DNA, with a velocity of roughly 100 bp/sec [22]; that experiment did not look at the translocation

velocity dependence on DNA tension but it is likely that the tension in the end-tethered DNAs used was less than 1 pN, making the result consistent with our model (Fig. 6b). Single-DNA experiments studying compaction against force see a stalling of compaction for DNA tensions in the ≈ 1 to 2 pN range [31, 34], with a distribution of steps in the 50-250 nm range, also consistent with the results of our model (Fig. 6d). Additional experiments have directly observed one-sided loop extrusion, with stalling observed for DNA tensions in the pN range [23]. Future experiments might try to disentangle DNA tension and load force dependence of translocation and loop extrusion, which show somewhat different behaviors in our model.

Direct evidence for directed translocation of condensin on DNA (chromatin) *in vivo* does not yet exist, but condensin-rich structures formed at the centromeres of chromosomes in budding yeast have been argued to come about as the result of translocation of condensin on DNA[55–57].

2. Eukaryotic cohesin.

Eukaryotic cohesin plays at least two distinct roles in genome organization. First, it acts to hold sister chromatids together during mitosis, and is degraded by a specific protease to allow sister chromatid segregation to occur. This activity is specific for cohesin and it is unclear whether it is at all related to SMC translocation or DNA loop extrusion discussed in this paper [58].

A second, more canonical function of cohesin is to compact and organize DNA during interphase and in mitosis and meiosis, presumably via DNA loop extrusion. Cohesin is involved in the formation of topologically isolated chromosomal domains (TADs), which are thought to control gene expression at long distances by restricting or promoting enhancer/promoter interactions. In vertebrates, the two boarders of a TAD are frequently defined by a pair of convergent CTCF binding sites. The precision of selection of orientation of CTCF sites at long distances observed in Hi-C experiments strongly suggests a mechanism based on processive 'tracking' or 'sliding' with CTCF acting as orientation-dependent barrier for cohesin translocation (Fig. 8c) [20, 21]. The loading of cohesin onto DNA appears to be somewhat more complicated than for condensin, and is dependent on well-characterized 'loading factors' which are necessary for ATP-dependent loading of cohesin onto DNA [36]. Unloading

of cohesin from DNA involves another protein, Wapl [59]

Once loaded, cohesin has been observed to diffusively slide on DNA *in vitro* [60, 61] but not to undergo clear directed translocations in the way recently observed for yeast condensin. Given the long distances apparently covered by cohesin during enhancer-promoter loop formation [20, 21] there is the question of whether diffusion along DNA can transport cohesin complexes sufficiently rapidly (*i.e.* so as to reach its destination in a single cell cycle). A key open question is whether cohesin moves along DNA itself (*e.g.*, via a motor mechanism similar to that discussed in this paper), whether whether it undergoes directed motion by being 'pushed' by other DNA-motile factors (e.g., RNApol or even condensin), or whether it simply diffuses.

In conclusion cohesin is a case of an SMC complex with factors regulating its loading and unloading (Fig. 8c), with the capacity to translocate rather large distances along DNA between those events. We note that the SMC1/3 heterodimer alone has been observed to be able to trap DNA loops of 150 nm size, of a specific chirality [33], as might be expected based on our loop-trapping translocator model (Fig. 5). If cohesin proves to not itself be a DNA motor, understanding why it does not have this function, while structurally similar condensin does, will be illuminating to understand.

3. Bacterial SMC complexes.

This paper has relied on the large amount of structural and functional data available for the bsSMC complex from *Bacillus subtilis* [24]. Single-molecule experiments on bsSMC [50] have indicated an ATP-stimulated (but ATP-hydrolysis-independent) compaction reaction, but have not yet observed step-like translocation. The single-DNA experiments also saw appreciable DNA compaction in the absence of ATP, perhaps suggesting that bsSMC may have a simple DNA-binding-compaction function. It might be that the DNA loop-capture step of Fig. 5 (or an alternate loop-capture function) may be able to proceed in some form.

In the cell, we know that most bsSMC is initially loaded near the chromosomal origin of replication, with the help of the ParB protein which associates with the *parS* DNA sequence [16, 17, 62]. It then appears to translocate along DNA, holding the two halves (the two 'replichores', Fig. 8d) of the circular chromosome together [16, 17], with a translocation velocity of approximately 1 kb/sec.

The high apparent velocity of bsSMC could be achieved by the mechanism of this paper by small-loop capture at about 3 cycles per second (100 nm contains 300 bp), or perhaps with less frequent transfer of compacted DNA structures (nucleoid-associated-protein-folded DNA or small supercoiled DNA structures) through the SMC upper compartment.

The MukBEF complex supports sister chromosome segregation in *E. coli*, possibly by establishing long-range DNA contacts as other SMC complexes do. There is a long history of observations of anucleate phenotypes associated with mutations in MukBEF [63], and recent *in vivo* observations suggest that the complex operates as a dimer of SMC complexes, and then forms clusters that compact and organize the bacterial chromosome [64]. As for eukaryote condensin, no specific loading factor for MukBEF has been identified (*E. coli* lack ParB and parS), implying that DNA loop extrusion would initiate at random positions. MukBEF complexes, however, localize in sub-cellular clusters, which are usually found near the replication origins [65]. Like CTCF in the case of cohesin, replication origins might act as barriers for MukBEF translocation and thus accumulate (Fig. 8e).

Single-molecule experiments on MukBEF have observed DNA compaction against forces of up to approximately 1 pN, but as observed for bsSMC, but without strong (on-off) ATP dependence [32], rather like what has been observed for bsSMC [50]. In addition, MukBEF has been observed to be able to "cross-link" separate double helices [66]. No experiment has yet directly observed clearly ATP-dependent DNA compaction, ATP-dependent translocation, or ATP-dependent loop extrusion by any bacterial SMC complex.

4. Eukaryotic SMC5/6.

The cellular functions of SMC5/6 are related to DNA double strand break repair, the resolution of sister DNA junctions during cell division and the control of DNA topology during DNA replication. The molecular mechanisms are poorly understood and the involvement of DNA loop extrusion is unclear. The structural similarities with other SMC complexes strongly suggest shared activities, and an ATP-dependent DNA-linking function has been observed *in vitro* [67]. Conceivably, the process of DNA loop extrusion might be used to identify and locate DNA junctions for repair and resolution (Fig. 8f).

Curiously, the SMC5 and Smc6 coiled coil arms are significantly shorter than the arms of other SMC proteins [42]. Our analysis predicts that the translocation by short-arm SMC

complexes might be somewhat more robust on stretched DNA. If so, then SMC5/6 complexes may have evolved to specifically deal with DNA under tension, potentially found near replication forks and during DNA repair. SMC5/6 structure may also reflect a preference for specific DNA topologies, e.g., crossings, loops or supercoiling of specific handedness.

E. Open questions

Many basic questions about SMC complexes remain unanswered. We have proposed a model for the translocation of SMC complexes on DNA, and it is by no means clear that all SMC complexes translocate on DNA. Possibly, translocation of some SMC complexes is regulated or driven by other proteins, as may be the case for cohesin. Single-molecule approaches can provide direct observation of translocation [22] and loop-extrusion [23] behavior, and it remains an open question how general this behavior is across different species of SMCs. For loop extrusion, it is possible that multiple translocators or extruders might "stack" to provide more robust motor function [3]. How extrusion and compaction forces scale with the number of extruders stacked is also a basic and open theoretical and experimental question. Finally, at the coarse-grained level of description of the models of this paper one is led to wonder what are the fundamental differences - if any - between the bacterial and eukaryote SMC complexes? For example, is there any role played by the apparent difference in symmetry between bacterial and eukaryote complexes (*e.g.*, homodimeric structure of 2x bsSMC vs. heterodimeric SMC2-SMC4 in condensin).

FUNDING

This work was supported by the National Institutes of Heath [grant numbers U54-CA193419 (CR-PS-OC), U54-DK107980, and R01-GM105847 to JM], by the European Research Council [grant number CoG 724482 to SG], by the Swiss National Science Foundation [grant number 200020_163042], and by the French Agence Nationale de la Recherche [grant number ANR-14-ACHN-0016 to AB].

Conflict of interest statement. None declared.

---

[1] Marko, J.F. (2009) Linking topology of tethered polymer rings with applications to chromosome segregation and estimation of the knotting length. *Phys. Rev. E* **79**, 051905.

[2] Marko, J.F. (2011) Scaling of Linking and Writhing Numbers for Spherically Confined and Topologically Equilibrated Flexible Polymers. *J. Stat. Phys.* **142**, 1353-1370.

[3] Alipour, E. and Marko, J.F. (2012) Self-organization of domain structures by DNA-loop-extruding enzymes. *Nucl. Acids Res.* **40**, 11202-11212.

[4] Schoeffler, A.J. and Berger, J.M. (2008) *Q. Rev. Biophys.* DNA topoisomerases: harnessing and constraining energy to govern chromosome topology. **41** 41-101.

[5] Adams, C. (2004) *The Knot Book*, AMS Press, Providence.

[6] Rybenkov, V.V., Ullsperger, C., Vologodskii, A.V. and Cozzarelli, N.R. (1997) Simplification of DNA topology below equilibrium values by type II topoisomerases. *Science* **277**, 690-693.

[7] Yan, J., Magnasco, M.O. and Marko, J.F. (1999) A kinetic proofreading mechanism for disentanglement of DNA by topoisomerases. *Nature*, **401** 932-935.

[8] Vologodskii, A. (2016) Disentangling DNA molecules. *Phys. Life Rev.* **18**, 118-134.

[9] Hirano, T. and Mitchison, T.J. (1994) A heterodimeric coiled-coil protein required for mitotic chromosome condensation in vitro. *Cell* **79**, 449-458.

[10] Uhlmann, F., (2016) SMC complexes: from DNA to chromosomes. *Nat. Rev. Mol. Cell Biol.* **17**, 399-412.

[11] Jeppsson, K., Kanno, T., Shirahige, K. and Sj ogren, C. (2014) The maintenance of chromosome structure: positioning and functioning of SMC complexes. *Nat. Struct. Mol. Cell Biol.* **15** 601-614.

[12] Rybenkov, V.V., Herrera, V., Petrushenko, Z.M. and Zhao, H. (2014) MukBEF, a chromosomal organizer. *J Mol Microbiol Biotechnol.* **24**, 371-383.

[13] Umbarger, M.A., Toro, E., Wright, M.A., Porreca, G.J., Bau, D., Hong, S.H., Fero, M.J., Zhu, L.J., Marti-Renom, M.A., McAdams, H.H., Shapiro, L., Dekker, J. and Church, G.M. The three-dimensional architecture of a bacterial genome and its alteration by genetic perturbation. (2011) *Mol. Cell* **44**, 252-264.

[14] Le, T.B., Imakaev, M.V. and Laub, M.T. (2013) High-resolution mapping of the spatial orga-

nization of a bacterial chromosome. *Science* **342**, 731-734.

[15] Marbouty, M., Le Gall, A., Cattoni, D.I., Cournac, A., Koh, A., Fiche, J.B., Mozziconacci, J., Murray, H., Koszul, R. and Nollman, M. (2015) Condensin- and Replication-Mediated Bacterial Chromosome Folding and Origin Condensation Revealed by Hi-C and Super-resolution Imaging. *Mol. Cell* **59**, 588-602.

[16] Wang, X., Le, T.B., Lajoie, B.R., Dekker, J., Laub, M.T. and Rudner, D.Z. (2015) Condensin promotes the juxtaposition of DNA flanking its loading site in Bacillus subtilis. *Genes Dev.* **29**, 1661-1675.

[17] Wang, X., Brandao, H.B., Le, T.B., Laub, M.T. and Rudner, D.Z. (2017) Bacillus subtilis SMC complexes juxtapose chromosome arms as they travel from origin to terminus. Science **355**, 524-527.

[18] Goloborodko, A., Marko, J.F. and Mirny, L.A. (2016) Chromosome Compaction by Active Loop Extrusion. *Biophys. J.* **110** 2162-2168.

[19] Goloborodko, A., Imakaev, M.V., Marko, J.F. and Mirny, L. (2016) Compaction and segregation of sister chromatids via active loop extrusion. *eLife* **5** e14864.

[20] Sanborn, A.L., Rao, S.S., Huang, S.C., Durand, N.C., Huntley, M.H., Jewett, A.I., Bochkov, I.D., Chinnappan, D., Cutkosky, A., Li, J., Geeting, K.P., Gnirke, A., Meinikov, A., McKenna, D., Stamenova, E.K., Lander, E.S. and Aiden, E.L. (2015) Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. *Proc. Natl. Acad. Sci. USA* **112** E6456-65.

[21] Fudenberg, G., Imakaev, M., Goloborodko, A., Abdennnur, N. and Mirny, L.A. (2016) Formation of Chromosomal Domains by Loop Extrusion. *Cell. Rep. 15* 2038-2049.

[22] Terekawa, T., Bisht, S., Eeftens, J., Dekker, C., Haering, C. and Greene, E. (2017) The condensin complex is a mechanochemical motor that translocates along DNA. *Science* **358**, 672-676.

[23] Ganji, M., Shaltiel, I.A., Bisht, S., Kim, E., Kalichava, A., Haering, C.H. and Dekker, C. (2018) Real-time imaging of DNA loop extrusion by condensin. *Science*, **360** 102-105.

[24] Diebold-Durand, M.-L., Lee, H., Ruiz Avila, L.B., Noh, H., Shin, H.-C., Im, H., Bock, F.P., B urmann, F., Durand, A., Basfeld, A., Ham, S., Basquin, J., Oh, B.-H. and Gruber, S. (2017) Structure of Full-Length SMC and Rearrangements Required for Chromosome Organization. *Mol. Cell* **67** 334-347.

[25] Seifert, F.U., Lammens, K., Stoehr, G., Kessler, B. and Hopfner, K.P. (2016) Structural mechanism of ATP-dependent DNA binding and DNA end bridging by eukaryotic Rad50. *EMBO J.* **35** 759-772.

[26] Liu, Y., Sung, S., Kim, Y., Li, F., Gwon, G., Jo, A., Kim, A.K., Kim, T., Song, O.K., Lee, S.E. and Cho, Y. (2016) ATP-dependent DNA binding, unwinding, and resection by the Mre11/Rad50 complex. *EMBO J.* **35** 743-758.

[27] Woo, J.S., Lim, J.H., Shin, H.C., Suh, M.K., Ku, B., Lee, K.H., Joo, K., Robinson, H., Lee, J., Park, S.Y., Ha, N.C. and Oh, B.H. (2009) Structural studies of a bacterial condensin complex reveal ATP-dependent disruption of intersubunit interactions. *Cell* **136**, 85-96.

[28] Murayama, Y. and Uhlmann, F. (2015) DNA Entry into and Exit out of the Cohesin Ring by an Interlocking Gate Mechanism. *Cell* **163**, 1628-1640.

[29] Piazza, I., Rutkowska, A., Ori, A., Walczak, M., Metz, J., Pelechano, V., Beck, M., Haering and C.H. (2014) Association of condensin with chromosomes depends on DNA binding by its HEAT-repeat subunits. *Nat. Struct. Mol. Biol.* **21**, 560-568.

[30] Zabrady, K., Adamus, M., Vondrova, L., Liao, C., Skoupilova, H., Novakova, M., Jurisinova, L., Alt, A., Oliver, A.W., Lehmann, A.R. and Palecek, J.J. (2016) Chromatin association of the SMC5/6 complex is dependent on binding of its NSE3 subunit to DNA. *Nucl. Acids Res.* **44**, 1064-1079.

[31] Strick, T., Kawaguchi, T. and Hirano, T. (2004) Real-time detection of single-molecule DNA compaction by condensin I. *Curr. Biol.* **14**, 874-880.

[32] Cui, Y., Petrushenko, Z.M. and Rybenkov, V.V. (2008) MukB acts as a macromolecular clamp in DNA condensation . *Nat. Struct. Mol. Biol.* **15**, 411-418.

[33] Sun, M., Nishino, T. and Marko, J.F. (2013) The SMC1-SMC3 cohesin heterodimer structures DNA through supercoiling-dependent loop formation *Nucl. Acids Res.* **41**, 6149-6160.

[34] Eeftens, J.M., Bisht, S., Kerssemakers, J., Haering, C.H. and Dekker, C. (2017) Realtime detection of condensindriven DNA compaction reveals a multistep binding mechanism *EMBO J.*, **36**, 3448-3457.

[35] Haering, C.H., Farcas, A.-M., Arumugam, P., Metson, J. and Nasmyth, K. (2008) The cohesin ring concatenates sister DNA molecules *Nature* **454**, 297-301.

[36] Murayama, Y. and Uhlmann, F., (2014) Biochemical reconstitution of topological DNA binding by the cohesin ring *Nature* **505**, 367-371.

[37] Cuylen, S., Metz, J. and Haering, C.H. (2011) Condensin structures chromosomal DNA through topological links *Nat. Struct. Mol. Biol.* **18**, 894-901.

[38] Wilhelm, L., B urmann, F., Minnen, A., Shin, H.C., Toseland, C.P., Oh, B.H. and Gruber, S. (2015) SMC condensin entraps chromosomal DNA by an ATP hydrolysis dependent loading mechanism in Bacillus subtilis. *eLife* **4**, e06659.

[39] Soh, Y.M., B urmann, F., Shin, H.C., Oda, T., Jin, K.S., Toseland, C.P., Kim, C., Lee, H., Kim, S.J., Kong, M.S., Durand-Diebold, M.L., Kim, Y.G., Kim, H.M., Lee, N.K., Sato, M., Oh, B.H. and Gruber, S. (2015) Molecular basis for SMC rod formation and its dissolution upon DNA binding. *Mol. Cell* **57** 290-303.

[40] Halford, S.E. and Marko, J.F. (2004) How do site-specific DNA-binding proteins find their targets? *Nucl. Acids. Res.* **32**, 3040-3052.

[41] Harkins, T.T. and Lindsley, J.E. (1998) Pre-steady-state analysis of ATP hydrolysis by Saccharomyces cerevisiae DNA topoisomerase II. 2. Kinetic mechanism for the sequential hydrolysis of two ATP. *Biochem.* **37** 7299-7312.

[42] B urrman, F., Basfeld, A., Vasquez Nunez, R., Diebold-Durand, M.L., Wilhelm, L. and Gruber, S. (2017) Tuned SMC Arms Drive Chromosomal Loading of Prokaryotic Condensin. *Mol. Cell* **65**, 861-872.

[43] van Noort, J., Verbrugge, S., Goosen, N., Dekker, C. and Dame, R.T. (2004) Dual architectural roles of HU: formation of flexible hinges and rigid filaments. *Proc. Natl. Acad. Sci. USA* **101**, 6969-6974.

[44] Skoko, D., Wong, B., Johnson, R.C. and Marko, J.F. (2004) Micromechanical analysis of the binding of DNA-bending proteins HMGB1, NHP6A, and HU reveals their ability to form highly stable DNA-protein complexes. *Biochem.* **43**, 13867-13874.

[45] Levene, S.D., Giovan, S.M., Hanke, A. and Shoura, M.J. (2013) The thermodynamics of DNA loop formation, from J to Z. *Biochem. Soc. Trans.* **41** 513-518.

[46] Sankararaman, S. and Marko, J.F. (2005) Formation of loops in DNA under tension. *Phys. Rev. E* **71**, 021911.

[47] Soga, N., Kimura, K., Kinosita Jr., K., Yoshida, M. and Suzuki, T. (2018) Perfect chemomechanical coupling of FoF1-ATP synthase. *Proc. Natl. Acad. Sci. USA* **114** 4960-4965.

[48] K. Nasmyth, (2001) Disseminating the genome: joining, resolving, and separating sister chromatids during mitosis and meiosis. *Annu. Rev. Genet.* **35** 673-745.

[49] Giuntoli, R.D., Linzer, N.B., Banigan, E.J., Sing, C.E., de la Cruz, M.O., Graham, J.S., Johnson, R.C. and Marko, J.F. (2015) DNA-Segment-Facilitated Dissociation of Fis and NHP6A from DNA Detected via Single-Molecule Mechanical Response. *J. Mol. Biol.* **427**, 3123-3136.

[50] Kim, H. and Loparo, J.J. (2016) Multistep assembly of DNA condensation clusters by SMC. *Nat. Comm.* **7**, 10200.

[51] Xiao, B., Johnson, R.C. and Marko, J.F. (2010) Modulation of HU-DNA interactions by salt concentration and applied force. Nucl. Acids Res. **38**, 6176-6185.

[52] Xiao, B., Zhang, H., Johnson, R.C. and Marko, J.F. (2011) Force-driven unbinding of proteins HU and Fis from DNA quantified using a thermodynamic Maxwell relation. Nucl. Acids Res. **39**, 5568-5577.

[53] Yan, J., Maresca, T.J., Skoko, D., Adams, C.D., Xiao, B., Christensen, M.O., Heald, R. and Marko, J.F. (2007) Micromanipulation studies of chromatin fibers in Xenopus egg extracts reveal ATP-dependent chromatin assembly dynamics. Mol. Biol. Cell 18, 464-474.

[54] Hirano, M. and Hirano, T. (1998) ATP-dependent aggregation of single-stranded DNA by a bacterial SMC homodimer. *EMBO J.* **17**, 7139-7148.

[55] Stephens, A.D., Haggerty, R.A., Vasquez, P.A., Vicci, L., Snyder, C.E., Shi, F., Quammen, C., Mullins, C., Haase, J., Taylor, R.M., Verdaasdonk, J.S., Falvo, M.R., Jin, Y., Forest, M.G. and Bloom, K. (2013) Pericentric chromatin loops function as a nonlinear spring in mitotic force balance. *J. Cell Biol.* **200** 757-772.

[56] Stephens, A.D., Quammen, C.W., Chang, B., Haase, J., Taylor, R.M. and Bloom, K. (2013) The spatial segregation of pericentric cohesin and condensin in the mitotic spindle. *Mol. Biol. Cell* **24**, 3909-3919.

[57] Stephens, A.D., Haase, J., Vicci, L., Taylor, R.M. and Bloom, K. (2011) Cohesin, condensin, and the intramolecular centromere loop together generate the mitotic chromatin spring. *J. Cell Biol.* **193**, 1167-1180.

[58] Gruber, S. (2017) Shaping chromosomes by DNA capture and release: gating the SMC rings. *Curr. Opin. Cell Biol.* **46** 87-93.

[59] Keung, S., Hegemann, B., Peters, B.H., Lipp, J.J., Schleiffer, A., Mechtler, K. and, Peters, J.M. (2006) Wapl controls the dynamic association of cohesin with chromatin. *Cell* **127**, 955-967.

[60] Stigler, J., Camdere, G.O., Koshland, D.E. and Greene, E.C. (2016) Single-Molecule Imaging

Reveals a Collapsed Conformational State for DNA-Bound Cohesin. *Cell Rep.* **15**, 988-998.

[61] Davidson, I.F., Goetz, D., Zaczek, M.P., Molodtsov, M.I., Huis In 't Veld, P.J., Weissman, F., Litos, G., Cisneros, D.A., Ocampo-Hafaila, M., Ladumer, R., Uhlmann, F., Vaziri, A. and Peters, J.M. (2016) Rapid movement and transcriptional re-localization of human cohesin on DNA. *EMBO J.* **35**, 2671.

[62] Minnen, A., B urmann, F., Wilhelm, L., Anchimiuk, A., Diebold-Durand, M.L. and Gruber, S. (2016) Control of Smc Coiled Coil Architecture by the ATPase Heads Facilitates Targeting to Chromosomal ParB/parS and Release onto Flanking DNA. *Cell Rep.* **14** 2003-2016.

[63] Niki, H., Jaffé, A., Imamura, R., Ogura, T. and Hiraga, S. (1991) The new gene mukB codes for a 177 kd protein with coiled-coil domains involved in chromosome partitioning of E. coli. *EMBO J.* **10**, 183-193.

[64] Badrinarayanan, A., Reyes-Lamothe, R., Uphoff, S., Leake, M.C. and Sherratt, D.J. (2012) In vivo architecture and action of bacterial structural maintenance of chromosome proteins. *Science* **338**, 528-531.

[65] Danilova, O., Reyes-Lamothe, R., Pinskaya, M., Sherratt, D. and Possoz, C. (2007) MukB colocalizes with the oriC region and is required for organization of the two Escherichia coli chromosome arms into separate cell halves. *Mol. Micro.* **65**, 1485-1492.

[66] Petrushenko, Z.M., Cui, Y., She, W. and Rybenkov, V.V. (2010) Mechanics of DNA bridging by bacterial condensin MukBEF in vitro and in singulo. *EMBO J.* **29**, 1126-1135.

[67] Kanno, T., Berta, D.G. and Sjogren, C. (2015) The Smc5/6 Complex is an ATP-Dependent Intermolecular DNA Linker. *Cell Rep.* **12**, 1471-1482.

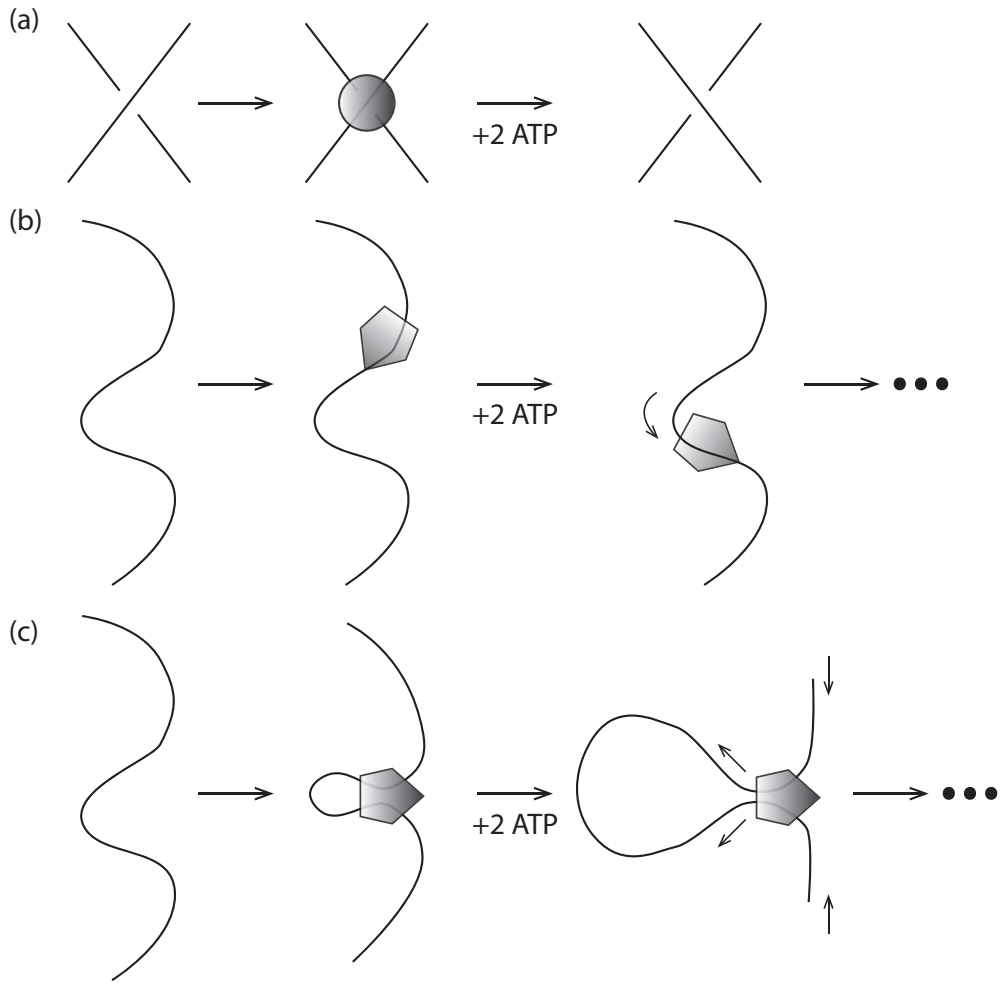


FIG. 1: Two active (ATP-hydrolysis-powered) protein machines responsible for DNA topology control *in vivo*.

(a) Type II topoisomerases are able to catalyze the transfer of a dsDNA segment through a second piece of dsDNA strand by generating a transient dsDNA break (not shown). Two ATPs are bound and hydrolysed per double-strand DNA exchange.

(b) SMC complexes are able to translocate along DNA. (c) SMC complexes are thought to be able to generate the gradual growth of size of DNA loops through active 'loop extrusion'. Note that only the first step of translocation and loop extrusion is shown; repetition of ATP binding and hydrolysis of additional ATPs can lead to processive enlargement of DNA loops.

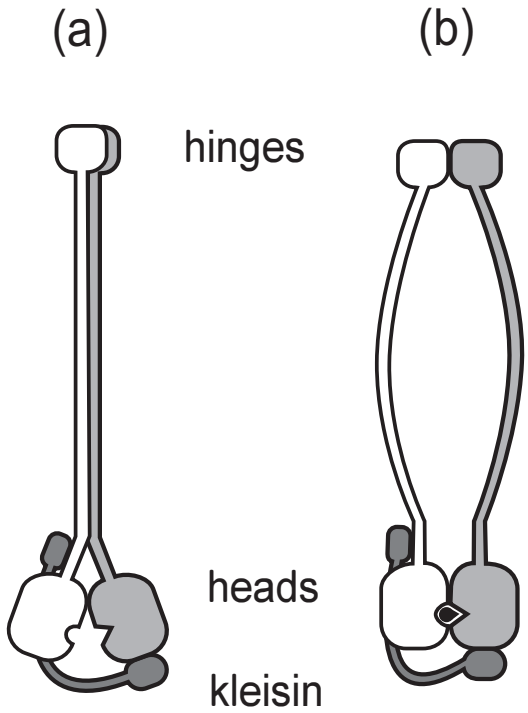


FIG. 2: Schematic view of SMC complex, showing the long (≈ 50 nm) coiled-coil domains which are linked at a 'hinge' domain and which have ATP-binding 'head' domains, and additional "kleisin" protein subunits which are able to bridge the head domains (dark gray). In the ATP-unbound (apo) state (a) the long coiled-coils associate (left). ATP (black) binding allows the head domains to bind together and change the conformation of the coiled-coils (b). ATP hydrolysis and release of hydrolysis products returns the enzyme to its apo state (a).

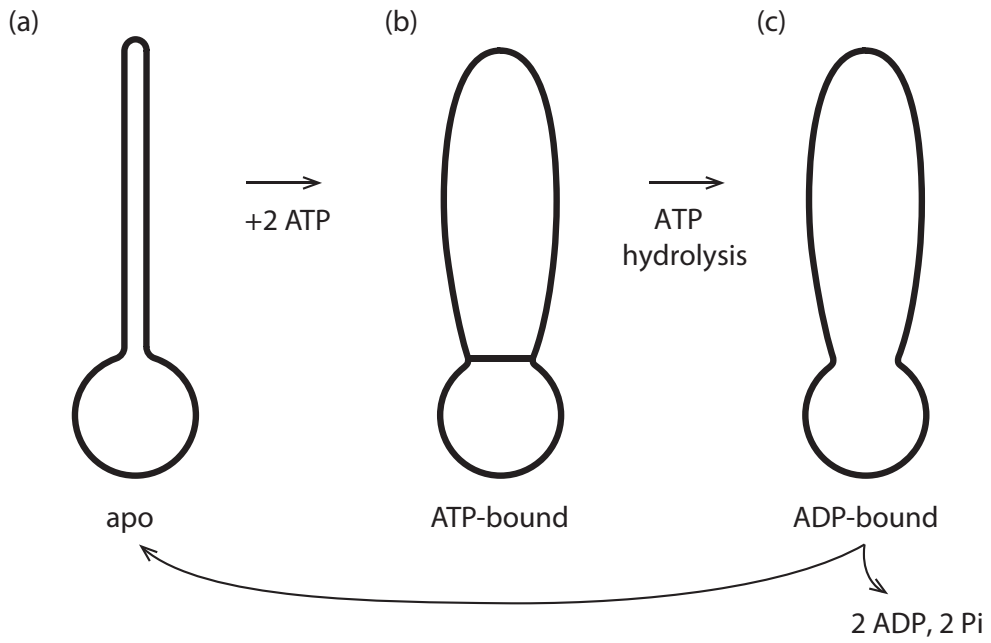


FIG. 3: SMC protein complex states.

(a) The apo (non-nucleotide-bound) state has a rod-shaped conformation with its two coiled-coil domains bound together and a small open 'lower' compartment;

(b) ATP binding opens the two coiled-coil domains resulting in the formation of a larger 'upper' compartment;

(c) ATP hydrolysis opens the 'gate' between the upper and lower compartment. Release of ADP leads to a reset of the enzyme to the apo state of (a).

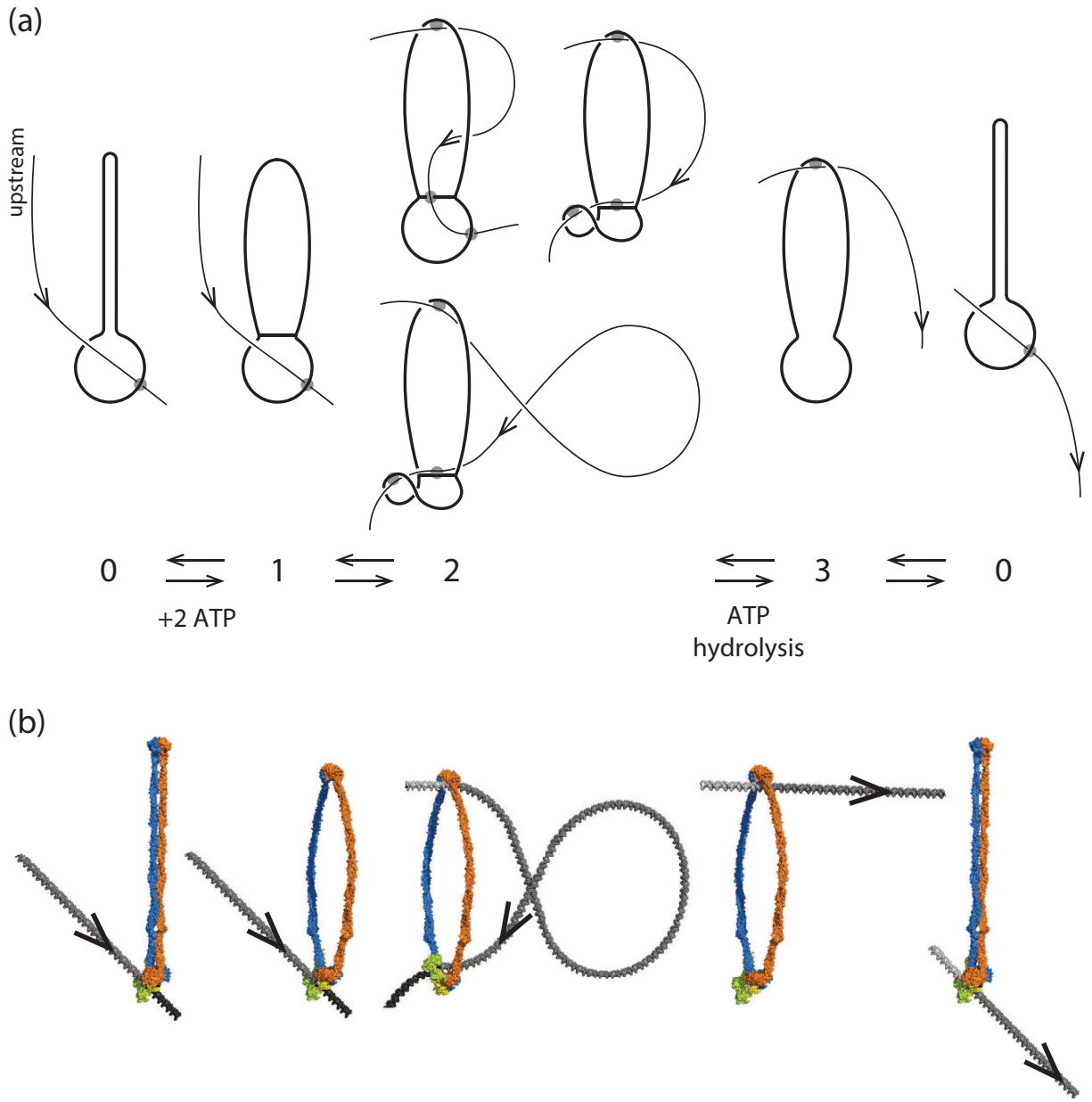
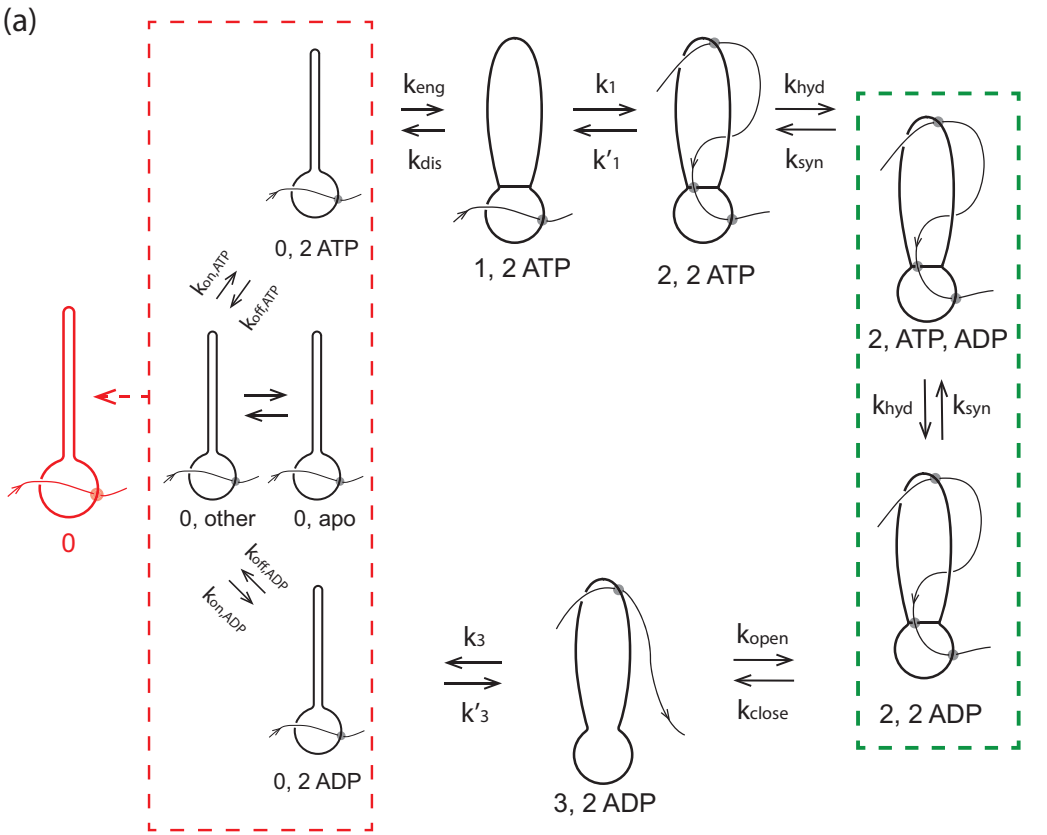

(a)
upstream
0
1
2
3
0
+2 ATP
ATP
hydrolysis
(b)

FIG. 4: Reaction diagram for SMC reaction with DNA that can generate translocation.

(a) Schematic/topological view of reaction, which starts with the SMC complex in the apo state 0 and with DNA bound in the lower compartment. Binding of ATP permits SMC head engagement and opening of the upper compartment (1), allowing a DNA loop to be captured in it (2). Then, ATP hydrolysis opens the wall between the two compartments, allowing the loop to pop out, transferring DNA contour length to the right (3). The coiled-coils then bind back together, pushing the DNA back into the lower compartment, returning the SMC-DNA complex to state (0), but translocated on the DNA by approximately the contour length captured in the loop in state (2). Three sketches of state 2 are shown, corresponding to a purely topological picture (state 2, top left), where the SMC complex lower chamber is bent to facilitate DNA loop capture (state 2, top right), and where a large DNA loop with a crossing is captured (state 2, bottom)

(b) Molecular structure diagram (based on data of Ref. [24]), showing corresponding states of SMC complex and DNA. Note the reorientation of the DNA segment by the "folding" of the lower compartment of the SMC complex in state 2, similar to the lower sketch of Fig. 4a. DNA structures with appropriate curvature were generated with the online tool 3D-DART and assembled together with manually fitted models of Smc-ScpAB in PyMol.


(a)
0, 2 ATP
keng
kdis
1, 2 ATP
k1
k′1
2, 2 ATP
khyd
ksyn
2, ATP, ADP
kon,ATP
koff,ATP
0
0, other
0, apo
khyd
ksyn
koff,ADP
kon,ADP
0, 2 ADP
k3
k′3
3, 2 ADP
kopen
kclose
2, 2 ADP

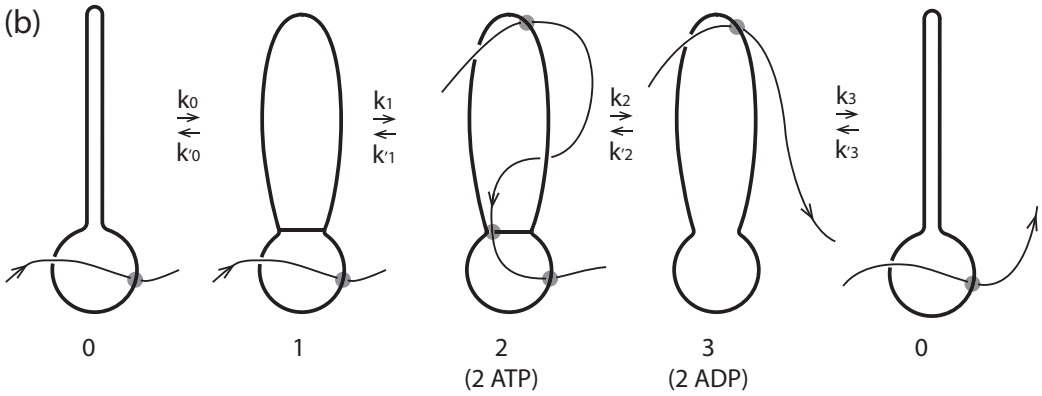

(b)
k0
k′0
k1
k′1
k2
k′2
k3
k′3
0
1
2
(2 ATP)
3
(2 ADP)
0

FIG. 5: Complete reaction diagram for SMC translocating along DNA.

(a) Complete reaction diagram starts with the SMC complex in state 0 and with DNA bound in the lower compartment, and able to bind nucleotides (ATP or ADP: left red dashed box includes 16 nucleotide-binding states; "other" represents the 13 states other than apo, 2 ATPs bound, and 2 ADPs bound). State labels are of the form N,n where N is the state number and n is bound nucleotide state (e.g., 0,2ATP refers to conformational state 0 with 2 ATPs bound). Binding of 2 ATPs permits SMC head engagement and opening of the upper compartment (state 1), allowing a DNA loop to be captured in it (state 2). ATP hydrolysis opens the wall between the two compartments, allowing the loop to pop out, transferring DNA contour length to the right (state 3). Finally the coiled-coils bind back together, pushing the DNA back into the lower compartment, returning the SMC-DNA complex to state 0, translocated on the DNA by approximately the contour length captured in the loop in state 2.

(b) Coarse-grained model, where the different nucleotide-binding states are combined together to form state 0 (left boxed states in (a)), and where ATP hydrolysis and loop release processes are combined together into a single 2-3 transition (the 2,ADP state of (a) is effectively "'integrated out"; state 2 now represents 2,ATP).

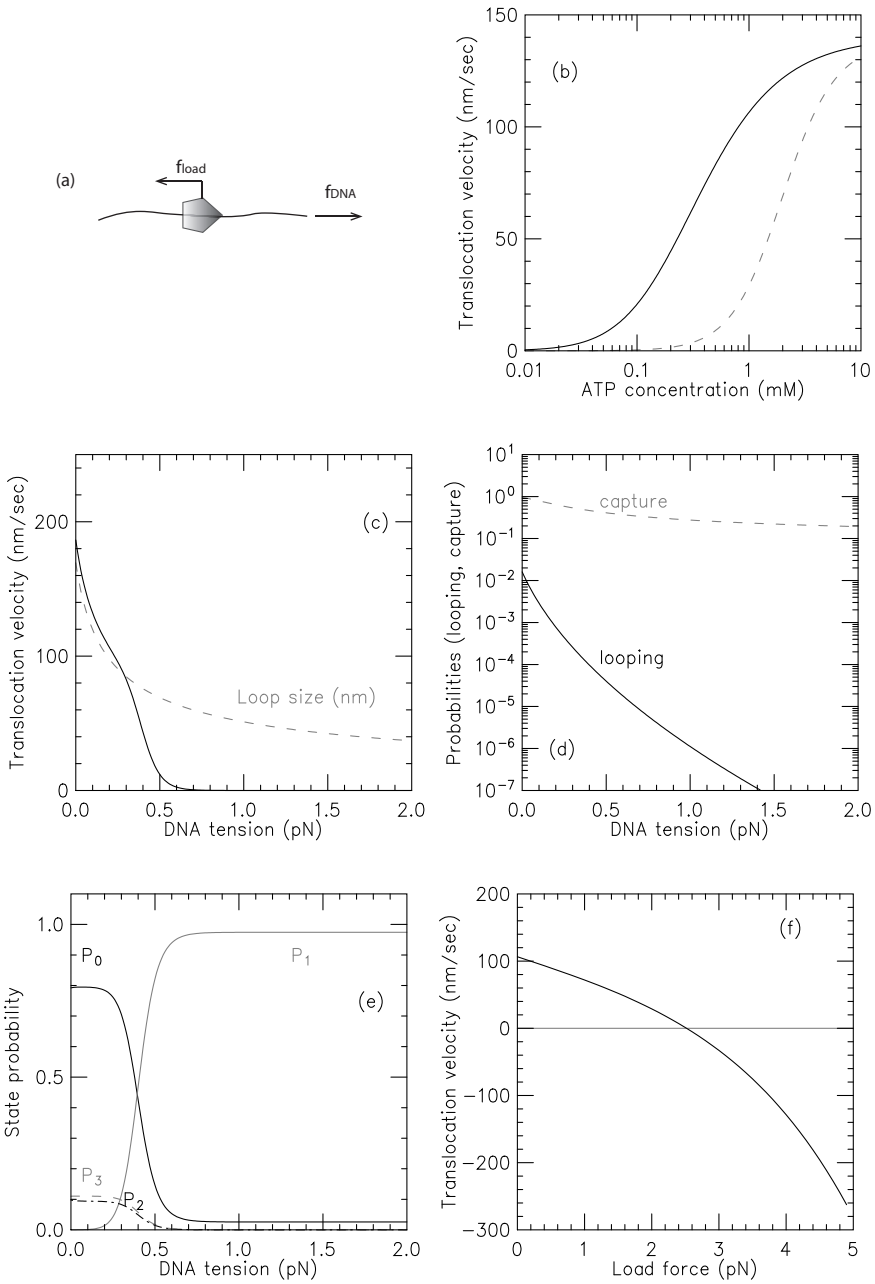

(a)
fload
fDNA
(b)
Translocation velocity (nm/sec)
ATP concentration (mM)
(c)
Translocation velocity (nm/sec)
Loop size (nm)
DNA tension (pN)
(d)
Probabilities (looping, capture)
capture
looping
DNA tension (pN)
(e)
State probability
P0
P1
P2
P3
DNA tension (pN)
(f)
Translocation velocity (nm/sec)
Load force (pN)

FIG. 6: Translocation for SMC DNA loop-capture model dependence on experimental control parameters.

(a) Distinction between tension in the DNA versus load force applied directly to the translocating SMC complex.

(b) Solid curve: velocity as a function of ATP concentration for zero load and DNA tension of $f_{\mathrm{DNA}} = 0.2$ pN. Dashed curve: result of same model in the presence of ADP; the velocity-ATP curve is simply shifted to larger ATP concentration values on the log [ATP] scale.

(c) Velocity as a function of DNA tension for zero load and ATP concentration of 1 mM shown by solid black curve; gray dashed curve shows the loop size. For the parameters chosen, DNA tension strongly (exponentially) suppresses SMC translocation for forces in excess of approximately 1 pN, but never reverses the direction of translocation.

(d) Variation of the loop-formation Boltzmann factor and the loop-capture probability in Eq. 4 with DNA tension, corresponding to (c).

(e) Probability of the four states of the model of Fig. 4, as a function of DNA tension, corresponding to (c).

(f) Velocity as a function of load force for template tension of 0.2 pN and ATP concentration of 1 mM. For the parameters chosen, the load force reverses the direction of the SMC complex along DNA for forces in excess of ≈ 3.5 pN.

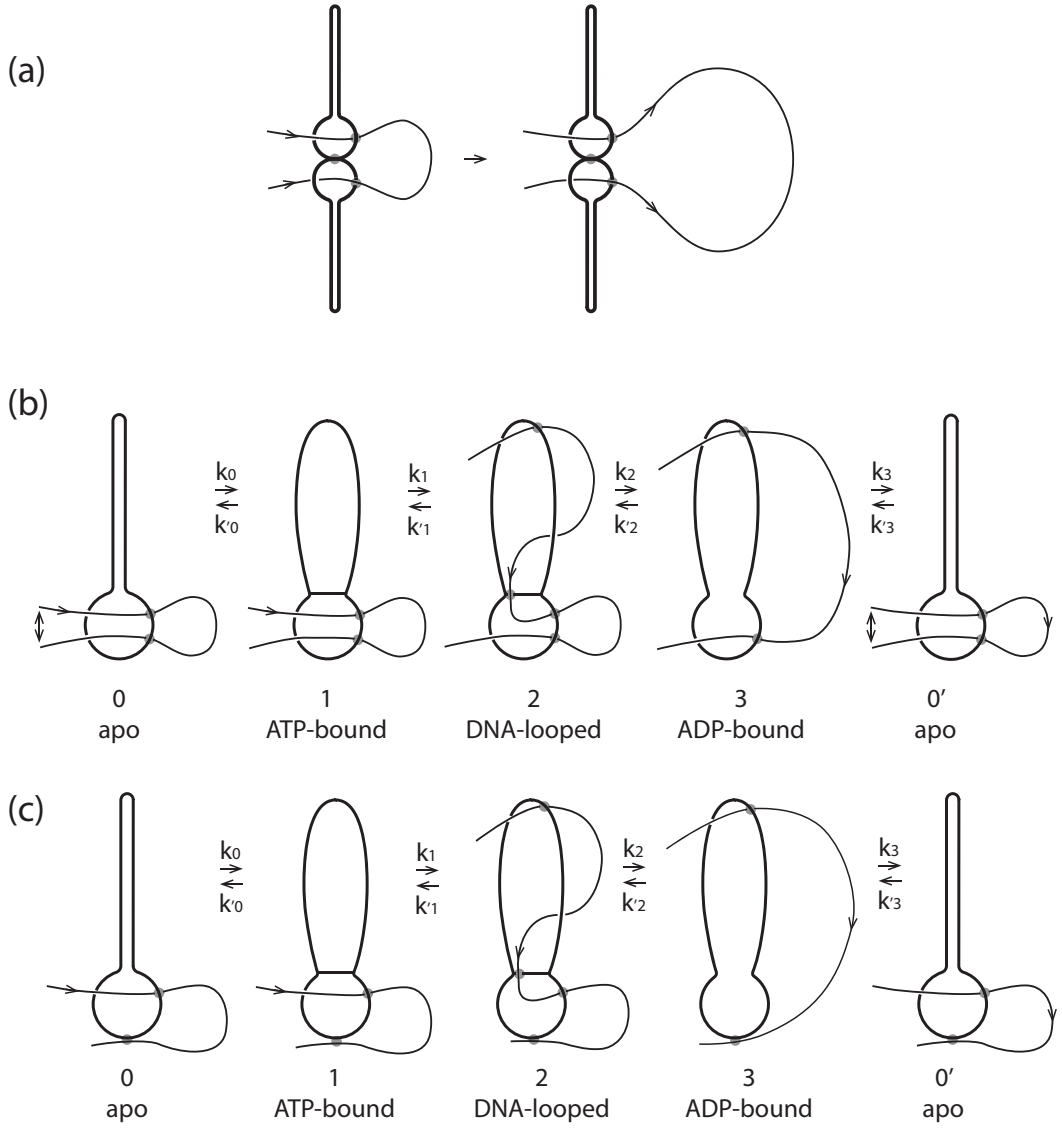


FIG. 7: Models for loop extrusion based on the translocator model of this paper.

(a) Two-translocator model. Two translocators are physically coupled together, and then loaded on DNA in orientations which generate loop growth.

(b) Single-translocator model. The reaction scheme is similar to that of Fig. 4, except that two DNAs can be trapped in the lower compartment.

(c) Asymmetric loop extrusion by a single translocator. The DNA remains bound to a spot outside the lower chamber (heavy black dot at bottom) resulting in one-sided loop extrusion.

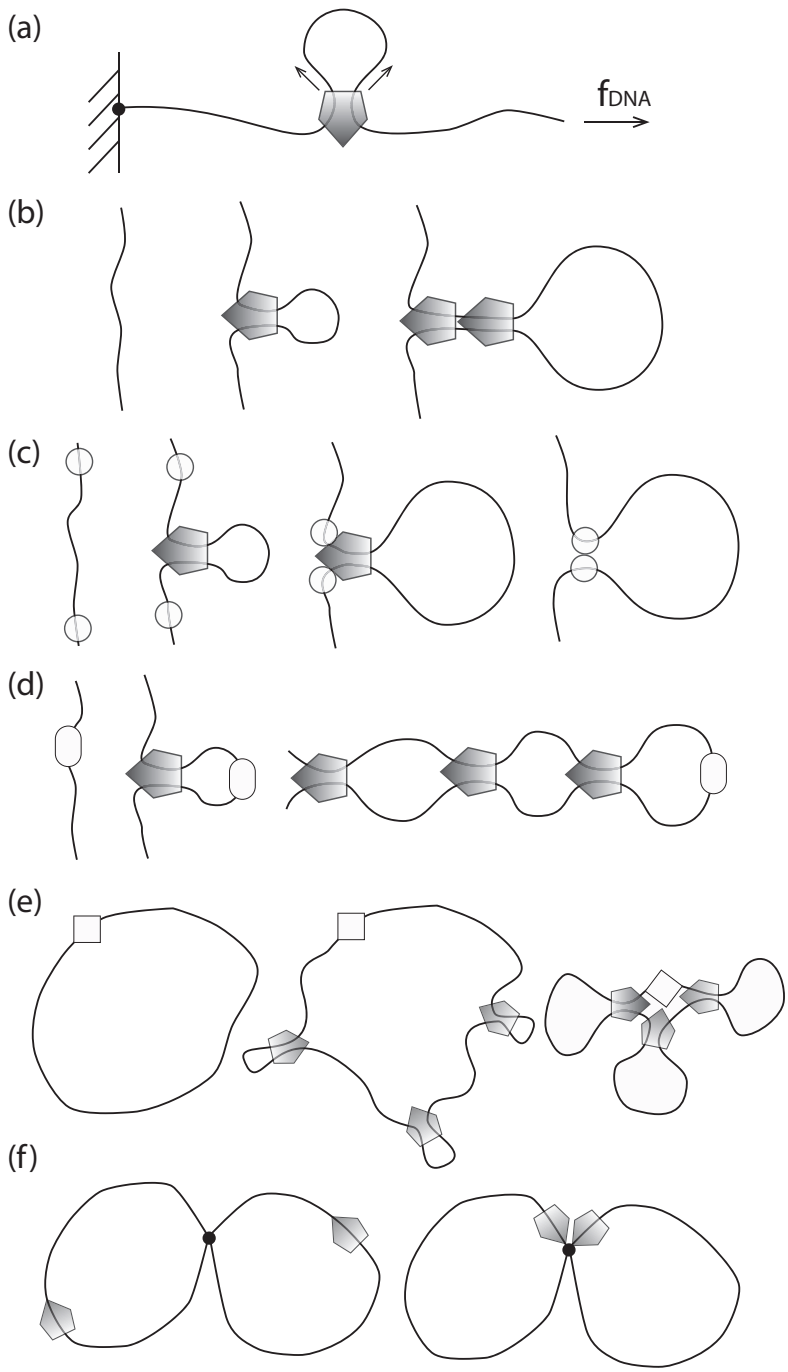

(a)
f_{DNA}
(b)
(c)
(d)
(e)
(f)

FIG. 8: Experimental single-molecule compaction and various modes of *in vivo* SMC complex regulation.

(a) Compaction assay with controlled DNA tension. The load force is related to applied tension f in a potentially complex way. The load force should be comparable but not larger than the tension.

(b) Eukaryotic condensin: nonspecific binding and stochastic unbinding with possible 'loop stacking'.

(c) Eukaryotic cohesin: nonspecific binding (using 'loading' proteins) with regulated halting/unbinding.

(d) Bacterial bsSMC from *B. subtilis*: site-specific binding (using the ParB protein bound to a *parS* site), with no regulation of unbinding so as to bind the two halves of a bacterial chromosome together.

(e) Bacterial MukBEF from *E. coli*: site-nonspecific binding with clustering at a specific site (the origin of replication).

(f) Eukaryotic SMC5/6: translocation of the complexes may lead to their co-localization at DNA junctions resulting from DNA recombination or incomplete DNA replication.