Biophysical Journal
Article

Biophysical Society

Ephemeral Protein Binding to DNA Shapes Stable Nuclear Bodies and Chromatin Domains

Chris A. Brackley,[1] Benno Liebchen,[1] Davide Michieletto,[1] Francois Mouvet,[1] Peter R. Cook,[2] and Davide Marenduzzo[1,*]
[1]SUPA, School of Physics and Astronomy, University of Edinburgh, Edinburgh, United Kingdom; and [2]Sir William Dunn School of Pathology, University of Oxford, Oxford, United Kingdom

ABSTRACT Fluorescence microscopy reveals that the contents of many (membrane-free) nuclear bodies exchange rapidly with the soluble pool while the underlying structure persists; such observations await a satisfactory biophysical explanation. To shed light on this, we perform large-scale Brownian dynamics simulations of a chromatin fiber interacting with an ensemble of (multivalent) DNA-binding proteins able to switch between an "on" (binding) and an "off" (nonbinding) state. This system provides a model for any DNA-binding protein that can be posttranslationally modified to change its affinity for DNA (e.g., through phosphorylation). Protein switching is a nonequilibrium process, and it leads to the formation of clusters of self-limiting size, where individual proteins in a cluster exchange with the soluble pool with kinetics similar to those seen in photobleaching experiments. This behavior contrasts sharply with that exhibited by nonswitching proteins, which are permanently in the on-state; when these bind to DNA nonspecifically, they form clusters that grow indefinitely in size. To explain these findings, we propose a mean-field theory from which we obtain a scaling relation between the typical cluster size and the protein switching rate. Protein switching also reshapes intrachromatin contacts to give networks resembling those seen in topologically associating domains, as switching markedly favors local (short-range) contacts over distant ones. Our results point to posttranslational modification of chromatin-bridging proteins as a generic mechanism driving the self-assembly of highly dynamic, nonequilibrium, protein clusters with the properties of nuclear bodies.

INTRODUCTION

In all living organisms, from bacteria to man, DNA and chromatin are invariably associated with binding proteins, which organize their structure (1–3). Many of these architectural proteins are molecular bridges that can bind at two or more distinct DNA sites to form loops. For example, bacterial DNA is looped and compacted by the histonelike protein H-NS, which has two distinct DNA-binding domains (4). In eukaryotes, complexes of transcription factors and RNA polymerases stabilize enhancer-promoter loops (5–8), while HP1 (9), histone H1 (10), and the polycomb-repressor complex PRC1/2 (11,12) organize inactive chromatin. Proteins also bind to specific DNA sequences to form larger structures, like nucleoli and the histone-locus, or Cajal and promyeloleukemia bodies (13–18). The selective binding of molecular bridges to active and inactive regions of chromatin has also been highlighted as one possible mechanism underlying the formation of topologically associated domains (TADs)—regions rich in local DNA interactions (6,8,19).

From a biophysical perspective, a system made up of DNA and DNA-binding proteins exhibits many kinds of interesting and seemingly counterintuitive behavior, such as the clustering of proteins in the absence of any attractive interaction between them. This process is driven by the "bridging-induced attraction" (20). In conjunction with the specific patterning of binding sites found in human chromosomes in vivo, this attraction can drive folding into TADs in the appropriate places on the chromosome (8).

In the simple case where there is only a nonspecific DNA-protein interaction (i.e., proteins can bind to any point on DNA), bridging-induced clustering can be understood as being due to a thermodynamic feedback loop: binding of bridges to multiple DNA segments causes an increase in local DNA concentration which, in turn, recruits further DNA-binding proteins, and further iterations then sustain the positive feedback. Subsequently, the ensuing clusters coarsen, and eventually phase-separate into one macroscopic cluster of DNA-bound bridges in equilibrium with

Submitted September 19, 2016, and accepted for publication January 6, 2017.

*Correspondence: dmarendu@ph.ed.ac.uk

Editor: Andrew Spakowitz.
http://dx.doi.org/10.1016/j.bpj.2017.01.025

a (diluted) pool of unbound proteins (21,22). In the more complex case with specific DNA-binding interactions, clustering is associated with the formation of DNA loops. Looped structures incur an entropic cost that increases superlinearly with the number of loops, and can stop the growth of a cluster beyond a critical size (8,23–26). Such specific binding drives the formation of promoter-enhancer loops (2); however there are several proteins that interact mainly nonspecifically with large regions of the genome, such as histone H1 and other heterochromatin-associated proteins (2). For this latter class of proteins, the abundance of binding sites in the nucleus would lead to clusters that coarsen progressively. However, this indefinite growth is not observed: we suggest that reversible posttranslational protein modifications may be the reason underlying the arrested coarsening.

Specifically, here we consider a biochemical reaction that can modify DNA-binding proteins. In our model, these proteins continuously switch between an active and an inactive state at rate α. While active, the proteins can thermodynamically bind to, and unbind from, DNA; but when inactive, proteins do not have any affinity for DNA. Such a reaction can arise in several biologically relevant scenarios. For instance, a complex of transcription factors and an RNA polymerase might stabilize a promoter-enhancer loop; upon transcription termination, the complex could dissociate and the loop disappear (2,3). Alternatively, phosphorylation, or other posttranslational modifications of transcription factors (27), may affect their affinity for chromatin, as might a conformational change in a protein or the reversible addition of a subunit to a protein complex, which might be driven by ATP hydrolysis.

In this work, we show that accounting for this nonequilibrium posttranslational mechanism strikingly broadens the range of physical behavior displayed by the chromatin/protein ensemble. In particular, we find that including switching proteins, which interact nonspecifically with a chromatin fiber, leads to qualitatively and quantitatively different results compared to nonswitching proteins. We observe that the continuous changing of state due to protein modification leads to an arrest of the coarsening of the clusters triggered by the bridging-induced attraction. Importantly, the size of the resulting clusters can be tuned by altering the switching rate α and we propose a mean-field calculation that supports this finding. Furthermore, we show that if proteins bind both specifically and nonspecifically, switching results in the formation of highly dynamic clusters, which are qualitatively different from those formed by nonswitching proteins. In the former case, proteins in the cluster exchange with the soluble pool, while the general shape of the cluster persists. These dynamic clusters recover quickly after simulated photobleaching, hence they closely resemble some of the nuclear bodies of eukaryotic cells. Finally, we consider a simplified model for the formation of TADs in chromosomes, and show that protein switching leaves the location of the domains unaltered, but strongly disfavors long-range inter-TAD interactions. All these findings point to an important and generic role of reversible protein modification in chromatin and nuclear organization.

MATERIALS AND METHODS

Details on the Brownian dynamics simulation method we use (including parameter values) and on the continuum mean field model (derivation, linear stability analysis, and amplitude equation) are given in the Supporting Material, where we also show additional results and figures that are discussed here in the main text.

RESULTS AND DISCUSSION

Protein switching arrests coarsening of chromatin bridges that bind nonspecifically

We perform Brownian dynamics simulations of a flexible chromatin fiber modeled as a bead-and-spring polymer (thickness 30 nm, persistence length 90 nm) interacting nonspecifically with either nonswitching or switching proteins. These proteins can bind to the fiber at more than one location (in our case through a Lennard-Jones potential; see the Supporting Material for details of the force field, and Fig. 1, *A* and *B*, for a schematic). For simplicity, we assume that the proteins are the same size as the chromatin beads (a realistic assumption, as each is likely to be a protein complex). We also assume proteins stochastically switch between on- and off-states at an equal rate, α. (Note: relaxation of either of these assumptions does not qualitatively alter our results.)

First, we consider the case of nonswitching proteins (i.e., we set $\alpha = 0$; Fig. 1 *C*). This case was previously studied in the literature (6,20,21,28), and it was shown to lead to polymer collapse (6,20,21,28) and clustering of proteins (20,21), depending on the protein concentration. For the concentrations used here, clusters coarsen and grow at the expense of smaller aggregates. During the early stages, this resembles the Ostwald ripening characteristic of liquid-gas phase separation; later on, we also observe coalescence of smaller clusters into larger ones (Movie S1). The average cluster size—measured as the number of bound proteins per cluster—increases with time with no sign of saturation until all clusters merge into one (Fig. 1 *Cii*; Movie S1). For early times, cluster size (which is also proportional to its volume) increases approximately linearly with time, as would be expected for Ostwald ripening in density-conserving model B (29). For later times, cluster growth is much slower, with a sublinear exponent (close to 0.25 for our parameters; Fig. 1 *Ei*). This slowing is due to the underlying polymer dynamics—as in blob formation during the collapse of a homopolymer, which is also slower than simple model B kinetics (30).

The dynamics with protein switching (i.e., $\alpha > 0$), is remarkably different: coarsening is completely arrested,

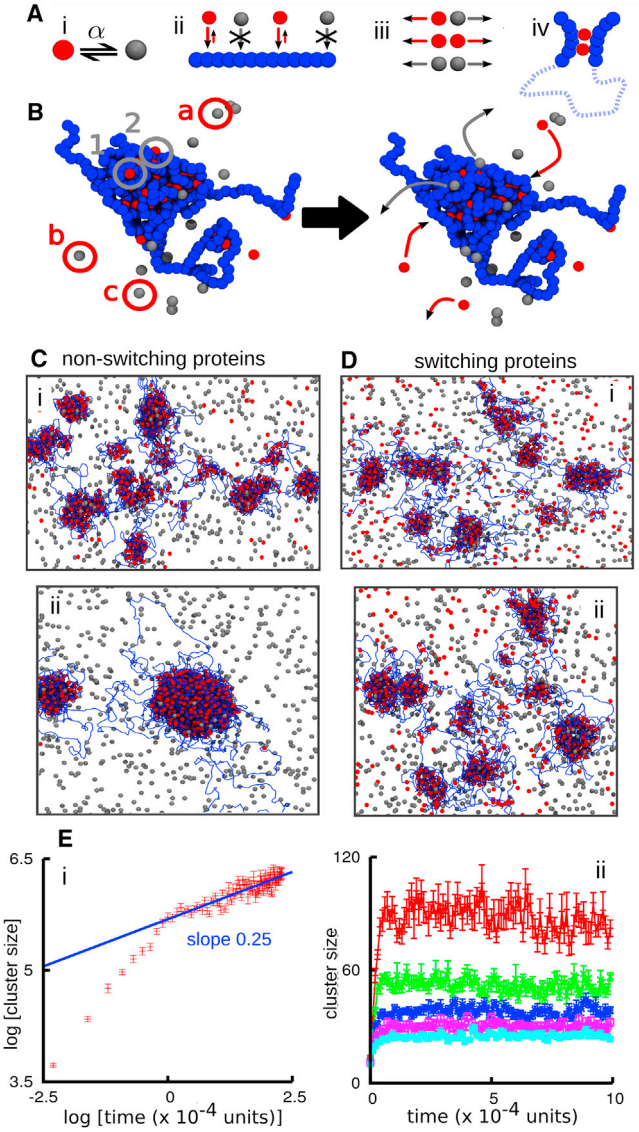

FIGURE 1 Protein switching arrests cluster coarsening. In (*A*)–(*D*), active and inactive proteins are colored red and gray, respectively; chromatin is represented by strings of blue beads. (*A*) Schematic of the model (Brownian dynamics simulations). (*i*) Proteins (*single spheres*) switch between red and gray states at rate α. (*ii*) Only proteins in the red state can bind chromatin. (*iii*) Red and gray beads interact via steric repulsion only. (*iv*) Proteins can bind to ≥ 2 sites to create molecular bridges and loops. (*B*) Snapshots illustrating protein binding/unbinding. Bound active proteins have clustered and compacted chromatin. Bound active proteins 1 and 2 (*gray circles*) switch and become inactive and dissociate (*gray arrows*); inactive proteins *a*–*c* in the soluble pool (*red circles*) are activated and may bind to the cluster (*red arrows*). (*C*) Snapshots taken (*i*) 10^4 and (*ii*) 2×10^4 simulation units after equilibration. The simulation involved a 5000-bead fiber (corresponding to 15 Mbp) and $N = 4000$ nonswitchable proteins, of which half are able to bind. (*D*) As in (*C*), but for $N = 4000$ switchable proteins ($\alpha = 0.0003$ inverse Brownian times). (*E*) Average cluster size as a function of time. Error bars denote standard deviations of the mean. (*i*) Nonswitching proteins. (*ii*) Switching proteins; from top to bottom, α equals 0.0001, 0.0002, 0.0003, 0.0004, and 0.0005 inverse Brownian times (or $\alpha^{-1} \simeq$ 10–60 s in real units). To see this figure in color, go online.

and the system achieves a microphase-separated state in which clusters have a well-defined average size (Fig. 1 *D*; Movie S2) that decreases with α (Fig. 1 *Eii*). The arrested phase separation can be understood intuitively as follows: on the one hand, thermodynamics dictates that the system should try to minimize interfaces, and this leads to coarsening, initially via Ostwald ripening, given the growth laws in Fig. 1 *Ei*; on the other hand, protein switching is a Poisson process, so active proteins switch off at a constant rate α, and leave the cluster. (Note: this is not the case for equilibrium proteins that can only unbind thermodynamically; not only is the unbinding rate slower, but such proteins are also highly likely to rebind to a nearby site before ever leaving a cluster.) Then, active proteins only have a timescale of the order α^{-1} in which to form a cluster before a significant proportion of proteins in that cluster inactivate. Hence, phase separation is arrested.

A mean field theory quantitatively explains the arrest of coarsening, and predicts average cluster size

To understand more quantitatively how protein switching arrests coarsening, we consider a simplified mean field theory that follows the time evolution of the chromatin density $\rho(\boldsymbol{x}, t)$, and the active protein density $\Phi(\boldsymbol{x}, t)$. Our equations describe the binding of the proteins to the chromatin together with the diffusion of all components, and they read as follows:

$$\begin{aligned} \dot{\rho} &= M_\rho \nabla^2 \left[a_1 \rho - k \nabla^2 \rho - \chi \Phi + g \rho^3 \right], \\ \dot{\Phi} &= M_\Phi \nabla^2 \left[a_2 \Phi - \chi \rho \right] - \alpha (\Phi - \Phi_0). \end{aligned} \tag{1}$$

These equations can be formally derived starting from a suitable underlying free energy density, and adding protein modification as a reaction term—the details are discussed in the Supporting Material. In the expressions in Eq. 1, M_ρ and M_Φ are the chromatin and protein mobility, respectively, so that $M_\rho a_1 \equiv D_1$ and $M_\Phi a_2 \equiv D_2$ represent effective diffusion coefficients, while χ is the coefficient describing bridging between active proteins and chromatin. Further, g captures steric repulsion in the chromatin fiber, k accounts for effective surface tension effects, and finally the last term in the equation for Φ models the biochemical reaction, where proteins switch from binding to nonbinding, and back, at a rate α. For $\alpha = 0$, the expressions in Eq. 1 ensure conservation of the global density of both chromatin and proteins—in other words, this is an example of generalized model B dynamics (17,29).

To identify the key parameters in our system, we now choose dimensionless time and space units $t_u = 1/\alpha$ and $x_u = \sqrt{D_2/\alpha}$ and redefine Φ as $\Phi(M_\rho \chi / D_2)$. In these units, our equations become

$$\dot{\rho} = \mathcal{D}_0 \nabla^2 \rho - A \nabla^4 \rho - \nabla^2 \phi + G \nabla^2 \rho^3, \tag{2}$$

$$\dot{\Phi} = \nabla^2\Phi - X\nabla^2\rho - (\Phi - \Phi_0), \quad (3)$$

so that the whole parameter space is spanned by the four dimensionless numbers $X = (\chi^2 M_\rho M_\Phi / D_2^2)$, $\mathcal{D}_0 = (D_1/D_2)$, $A = \alpha k M_\rho/(D_2^2)$, and $G = gM_\rho/D_2$.

One solution of Eqs. 2 and 3 is given by the uniform phase $(\rho, \Phi) = (\rho_0, \Phi_0)$, which is stable in the absence of bridging ($\chi = 0$). To see how the interplay of bridging and biochemical switching can create patterns, we performed a linear stability analysis of this uniform state (Fig. 2, detailed in the Supporting Material). The result is that small perturbations of the uniform phase grow if $X > X_c = (\sqrt{A} + \sqrt{\mathcal{D}})^2$, where $\mathcal{D} = \mathcal{D}_0 + 3G\rho_0^2$. This instability criterion translates in physical units to

$$\chi > \sqrt{k\alpha/M_\Phi} + \sqrt{(D_2/M_\Phi)[D_1/M_\rho + 3g\rho_0^2]}. \quad (4)$$

From Eq. 4, it is important to notice that the instability toward nonuniform phases is driven by bridging (i.e., χ needs to exceed a certain threshold), whereas diffusion of chromatin D_1 and proteins D_2, the excluded volume g, and the protein modification α, all tend to stabilize uniform chromatin-protein distributions. This can be readily understood by considering rapidly switching proteins (i.e., the case of large α). In this scenario, proteins switch between the on- and off-states so rapidly that cluster formation becomes possible only by increasing the binding strength χ. A similar argument applies if the diffusion of the components is very fast or the short-range repulsion very strong. Finally, Eq. 4 also tells us that there is no lower bound in protein concentration for the onset of this bridging-induced instability, because Φ_0 does not appear in the equation.

Calculating the wavenumber at the onset of instability (see the Supporting Material) unveils the remarkable role played by the biochemical reaction for structure formation. Specifically, we find $q_c = (\mathcal{D}/A)^{1/4}$ for the dimensionless onset wavenumber, translating in physical units to the typical length scale of

$$L_c = 2\pi\left[\frac{D_1 D_2 + 3M_\rho g\rho_0^2 D_2}{\alpha k M_\rho}\right]^{1/4}. \quad (5)$$

Hence, in contrast to models without protein modification, this system exhibits a short wavelength instability (Fig. 2 *A*), which turns into a long wavelength instability only in the limit $\alpha \to 0$ (which would lead to $L_c \to \infty$, *dotted black line* in Fig. 2 *A*). Our linear stability analysis therefore suggests that the presence of the biochemical reaction has qualitative consequences for the clustering in the system, in that it leads to self-limiting cluster sizes, or put differently, to microphase separation rather than to macrophase separation—in full agreement with the simulations shown in Fig. 1.

To further confirm that within our mean-field theory, clusters cannot coarsen indefinitely, we also performed a weakly nonlinear expansion, through which we found that the amplitude of the chromatin density fluctuations close to the uniform state obey the "real Ginzburg Landau equation" (which is associated with formation of stationary patterns of well-defined self-limiting size (31); and also see the Supporting Material). Finally, Eq. 5 also predicts that, at least close to the onset of clustering, the average number of proteins in any aggregate should scale as $L_c^3 \sim \alpha^{-3/4}$. This behavior is reported in Fig. 2 *B*, where we show the very good agreement of this scaling argument with the results from the Brownian dynamics simulations.

While estimating values for the parameters appearing in Eq. 5 is challenging, the cluster size predicted by our model through the Brownian dynamics simulations compares favorably with that obtained experimentally. In Fig. 2 *B*, the typical number of proteins in a cluster found by simulations ranges from 25 to 100 for $\alpha \simeq 10^{-4} - 10^{-3}$ inverse Brownian times. From these values, one can extract the typical cluster size as 70–100 nm (a protein, here, is ~30 nm). These values are in line with, for example, STORM experiments performed on Polycomb nuclear bodies in *Drosophila* (11,12). As the Brownian time is $\tau_B \simeq 6$ ms (see the Supporting Material), the inverse switching rate is ~1 min, which is reasonable for post-translational modifications such as phosphorylation or acetylation (32).

This specific example shows that the cluster size we get in our simulations is similar to that of nuclear protein clusters. We will come back to a comparison to experiments in the next section, where we consider the case of specific binding, which is more relevant to nuclear bodies in vivo. There, we also simulate a typical photobleaching experiment to assess the kinetic recycling of nuclear bodies.

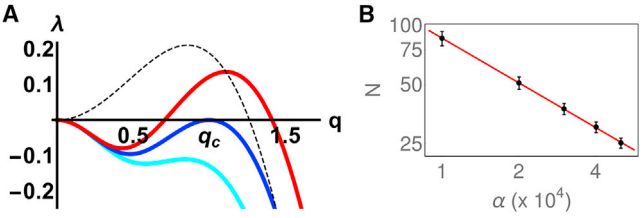

FIGURE 2 Mean field theory predicts arrested coarsening with protein modification. (*A*) Dispersion relation, showing the growth rate, λ, as a function of the magnitude of the wavevector, q, for fluctuations around the uniform solution of Eq. 2, for $\mathcal{D} = A = 1$, and $X = 3.5$ (*cyan*), corresponding to linear stability of the uniform phase, $X = X_c = 4.0$ (*blue*), marking the onset of instability, and $X = 4.5$ (*red*), revealing the growth of clusters with a characteristic length scale. (*Dotted black line*) Typical dispersion relation in the absence of protein modification, which leads to a long wavelength instability. (*B*) Scaling between number of proteins in a cluster and switching rate found from Brownian dynamics simulations. Points show saturation values (±SD) of number of particles per cluster N (after 1.5×10^5 simulation units); the line shows a least-squares fit with a slope of −0.756. To see this figure in color, go online.

Switching proteins with specific binding self-assemble into recycling nuclear bodies

The model considered in Fig. 1 assumes that proteins bind nonspecifically. While this is a good approximation for generic heterochromatin-binding proteins in silenced regions of the genome, most transcription factors bind strongly to specific sites in active regions and to most other DNA nonspecifically (33). Therefore, we consider proteins binding with high affinity to every 20th bead (i.e., every 60 kbp), and with low affinity to all others. (Note: similar results are expected for different patterns of binding sites (8,20).) Now bound proteins self-assemble into clusters of self-limiting size even when $\alpha = 0$ (Fig. 3; Movie S3). In other words, coarsening is always arrested. As suggested previously (20,23,25), specific binding creates loops that are associated with entropic costs that scale superlinearly with loop number. In turn, this limits cluster growth (23).

Although coarsening is arrested whatever the value of α, there is still a major difference between the dynamics of the equilibrium and switching proteins. Without switching, proteins can only unbind thermodynamically, which requires a long time: as a result, proteins rarely exchange between clusters (Fig. 3; Movie S3). With switching, there is a constant turnover of proteins within the clusters, which recycle all their components over a time $\sim\alpha^{-1}$ (Fig. 3; Movie S4). Reducing the strength of the specific interactions can also lead to protein turnover (Fig. S1), but this requires fine-tuning of the parameters to simultaneously ensure stable binding and the recycling of proteins in clusters. In contrast, protein modification naturally leads to such recycling for any values of specific and nonspecific binding affinity.

To quantitatively characterize the dynamics of turnover within clusters, we perform a simulated fluorescence-recovery-after-photobleaching (FRAP) experiment (34). In such an experiment some of the clusters are photobleached at a given time, and recovery of fluorescence is then monitored (Fig. 3). The fluorescence signal (proportional to the number of nonphotobleached active proteins in the clusters) recovers quickly in the $\alpha > 0$ case (Fig. 3, *Aiii* and *B*), but not in the $\alpha = 0$ case (Fig. 3, *Aiv* and *Bi*), at least for large values of the specific interaction strength. The dynamics of recovery can be measured by counting the number of unbleached proteins in the photobleached volume (Fig. 3 *Bi*); this is proportional to the fluorescence intensity measured in a standard FRAP experiment. Alternatively, the number of unbleached proteins in clusters can be used (Fig. 3 *Bii*). Both approaches give similar recovery timescales, and confirm that protein modification is required to create clusters in which proteins can recycle.

The clusters found in Fig. 3 typically contain ~20–100 proteins that recycle (Fig. S2 *A*) and give average cluster sizes of ~70–100 nm. Cluster size depends on both protein concentration and interaction energy (e.g., in Fig. S2 *B*, there are only ~5–10 proteins per cluster). Therefore, this mechanism can produce clusters with a wide range of sizes. Note that nuclear

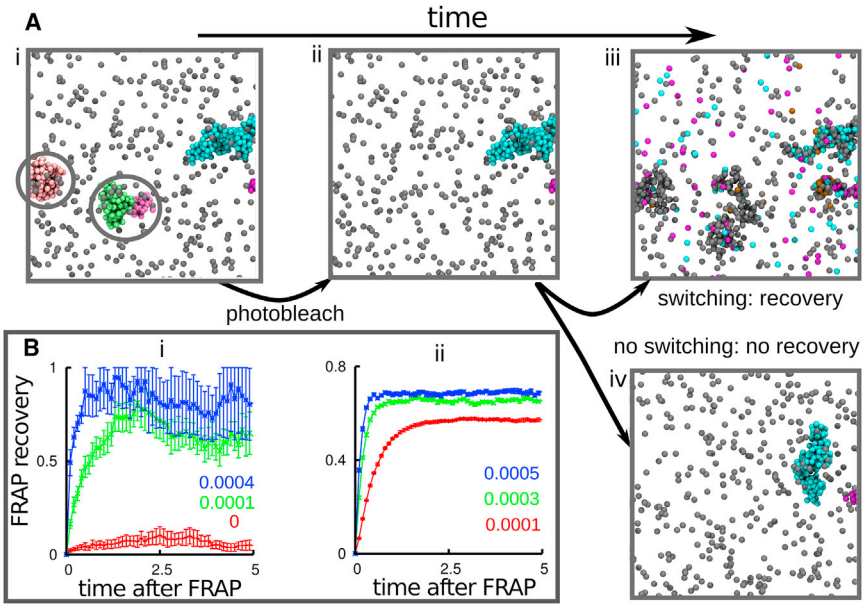

FIGURE 3 In silico FRAP (Brownian dynamics simulations). (*A*) Snapshots taken 10^4 (*i* and *ii*) or 2×10^4 (*iii* and *iv*) after equilibration, during an in silico FRAP experiment (only proteins—and not chromatin beads—are shown for clarity). (*i*) The simulation begins with $N = 2000$ equilibrium proteins, half of which are able to bind the chromatin fiber, both specifically (interaction strength 15 k_BT, cutoff 1.8σ) to every 20th bead in the polymer, and nonspecifically (interaction strength 4 k_BT, cutoff 1.8σ) to any other bead. After 10^4 time units, a structure with multiple clusters forms. The snapshot shows only a portion of this, for clarity; five clusters of bound proteins have developed (unbound proteins are *gray*; bound proteins in the five clusters are *blue*, *pink*, *purple*, and *green*). Circled areas will be photobleached. (*ii*) Photobleaching involves making bound proteins invisible (the bleached proteins are still present in the simulation). (*iii*) If proteins can switch, clusters reappear in the same general place (as new proteins replace their bleached counterparts). (*iv*) If proteins cannot switch (i.e., $\alpha = 0$), clusters do not recover (as their protein constituents do not recycle). (*B*) FRAP recovery. Error bars give SD of mean, and time is given in multiples of 10^4 simulation units; the values of α, in units of inverse Brownian times, are as indicated in each panel. Only the postbleaching signal is shown (the prebleaching value would be constant and equal to 1 in these units). (*i*) Number of unbleached proteins in the bleached volume (a sphere of 50σ) as a function of time, after bleaching. The signal is normalized with respect to the number of proteins initially in the bleached volume. (*ii*) Number of unbleached proteins in clusters as a function of time after bleaching, after all proteins in clusters at a given instant are bleached. The signal is normalized with respect to the proteins in clusters at the time of bleaching. To see this figure in color, go online.

bodies range from large nucleoli (up to several micrometers), through Cajal and promyelocytic leukemia bodies (~1 μm) (14–16), to transcription factories containing ~10 active transcription complexes, and to polycomb bodies (~100 nm) (3,11,35,36). Importantly, like most nuclear bodies, our clusters also retain a memory of their shape. Thus, in Fig. 3 *A*, when most of the components of the pink cluster on the left have turned over, the general shape of the cluster persists (see also Movies S4, S5, and Fig. S3). This is because the chromatin scaffold associated with the protein clusters (i.e., the sites of specific binding) retains a general three-dimensional structure that does not change much over time (Fig. S4). Taken together, these results strongly support the conjecture that nuclear bodies emerge from the aggregation of bound switching proteins, and that switching both arrests phase separation and ensures that bound proteins continually exchange with the soluble pool.

Notably, the nuclear bodies that our clusters resemble generally show FRAP recovery times in the range of tens of minutes (37–39). These are too slow to be accounted for by diffusion, and too fast to be compatible with the thermodynamic unbinding of tightly bound proteins (see the Supporting Material); remarkably, our simulations can instead readily account for these timescales.

Thus, within our model, the recovery time over which nuclear bodies recycle their proteins is linked to protein modification, and it is simply proportional to α^{-1}. Typical rates of posttranslational protein modification can be of the order of minutes (and will be slower within nuclear bodies due to macromolecular crowding), and transcription termination also occurs within minutes after initiation. In light of this, our simulations predict recovery timescales of the order of α^{-1}, or minutes, in broad agreement with those measured experimentally (37–39). Further to this, there is biological evidence that protein modifications can take place within nuclear bodies (37). For instance, enzymes performing posttranslational modifications are found in Cajal bodies (37), and phosphorylation or ubiquination of the BMI1 subunit of the PcG PRC1 complex are important factors that determine the kinetics of exchange in polycomb bodies (38).

Protein switching preserves TAD structure, while suppressing long-range interactions

Clustering of bridging proteins can lead to the formation of chromatin interaction domains (6,8,12,20) resembling TADs found in Hi-C data (19). It is therefore of interest to ask how switching affects TAD structure and dynamics. Here, we return to a toy model first considered elsewhere (8): the fiber has a regular pattern of binding and nonbinding regions (Fig. 4 *A*), and each binding region spontaneously and reproducibly assembles into a TAD that is flanked by

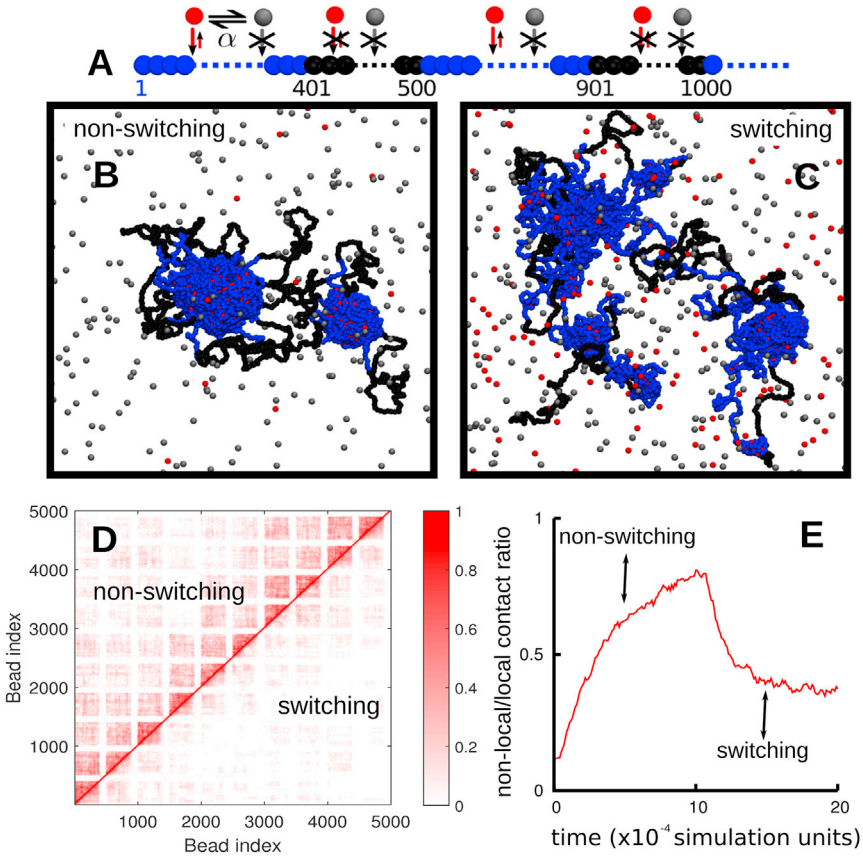

FIGURE 4 Switching promotes intra-TAD contacts, but suppresses inter-TAD ones. (*A*) Overview: Simulations involved $N = 2000$ nonswitching ($\alpha = 0$) or $N = 2000$ switching proteins ($\alpha = 0.0001$ inverse Brownian times); for $\alpha = 0$, half of the proteins are binding. In both cases, interaction energy and cutoff values were 4 k_BT and 1.8σ. The fiber (length 15 Mbp) consisted of regularly interspersed segments containing runs of binding (*blue*) and nonbinding (*black*) beads (segment sizes 1.2 Mbp and 300 kbp, respectively). (*B* and *C*) Snapshots taken after 10^5 simulation units. Nonbinding (*gray*) and binding (*red*) proteins are shown. (*D*) Contact maps (averages from 10 simulations) for nonswitching (*top-left triangle*), and switching proteins (*lower-right triangle*). The scale (*right*) indicates contact frequencies. (*E*) The evolution of the ratio of nonlocal contacts over time. A local (nonlocal) contact is one between beads separated by less (more) than 1.2 Mbp along the fiber. Here the simulation was run for 10^5 simulation units with nonswitching proteins; switching was then turned on ($\alpha = 0.0001$ inverse Brownian times) and the simulation was run for a further 10^5 simulation units. To see this figure in color, go online.

a disordered nonbinding region (8). The regular interspersion of nonbinding segments in Fig. 4 *A* fixes the locations of TAD boundaries; consequently, clusters form (Fig. 4, *B* and *C*) at reproducible positions along the fiber, and this in turn yields TADs seen in averaged contact maps (Fig. 4 *D*). Such patterns resemble those seen in Hi-C data obtained from cell populations.

Variations in α have several effects (Fig. 4). First, the configurations found at steady state are qualitatively different. Although cluster growth is limited for both $\alpha = 0$ and $\alpha > 0$, the (recycling) clusters formed by switching proteins are much smaller (Fig. 4, *B* and *C*; Movie S6). Second (and notwithstanding this qualitative difference), the contact maps close to the diagonal are remarkably similar (Fig. 4 *D*; compare patterns on each side immediately next to the diagonal); this indicates that local TAD structure is largely unperturbed by switching. However, for $\alpha > 0$, nonlocal contacts (i.e., between chromatin segments far apart along the fiber) are strikingly suppressed (Fig. 4 *D*, compare patterns on each side far from the diagonal, see also Fig. S5), and higher-order folding of one TAD onto another is suppressed.

This observation can be explained as follows. First, the timescale for the formation of TADs is comparable to (or smaller than) that of protein recycling within a TAD (see Supporting Material for an estimate of such timescales). Computer simulations of TAD formation in *Drosophila* and human chromosomes also suggest that the local structure can be formed very rapidly (at most, in minutes) (8,12). Therefore, it is plausible that local TAD folding is fast enough not to be perturbed much by protein modification. Second, when a particular protein switches from binding to nonbinding, a contact is lost, and it is likely that local ones can reform faster than nonlocal ones.

In light of this, the nonequilibrium switching we account for in this model provides a mechanism allowing faster large-scale rearrangements, and a more effective trimming of entropically unfavorable long-ranged interactions. In other words, active posttranslational modification tilts the balance in favor of local intra-TAD contacts at the expenses of inter-TAD ones. This observation is consistent with the sharp decay beyond the Mbp scale seen in Hi-C data (19,40).

CONCLUSIONS

We have shown that active posttranslational protein modication (e.g., phosphorylation, methylation, acetylation (27), or any other nonequilibrium reaction where a protein switches between active and inactive states) has a profound and generic effect on the behavior of a chromatin-protein mixture. Our results strongly suggest that the interplay between protein bridging and protein modification is an important principle underlying nuclear organization within eukaryotes.

First, it was previously shown that nonswitching proteins able to bind nonspecifically to chromatin to form molecular bridges assemble into clusters which have a natural tendency to coarsen (20,21). In this work, we have shown that accounting for protein switching changes the qualitative and quantitative behavior of the system; cluster growth is self-limiting (Fig. 1)—a phenomenon that can be understood via a simple mean field theory (Fig. 2). This theory also provides an example of arrested phase separation, and it can explain why nuclear bodies do not progressively grow (13–16), and why neighboring clumps of heterochromatin—whether detected using classical staining and microscopy, or through inspection of Hi-C contact maps (19,41)—rarely merge into one superdomain.

Second, nonswitching proteins bind specifically to cognates sites on the chromatin fiber, and they also cluster; however, specific binding is known to arrest the coarsening through the entropic penalty of loop formation (8,20). But in contrast to what is seen in photobleaching experiments (37,38), bound proteins in the ensuing clusters exchange little with the soluble pool. Moreover, the timescales seen in such bleaching experiments are too slow to be accounted for by diffusion, and too fast to be compatible with the thermodynamic unbinding of tightly bound proteins. The results reported in this work strongly suggest that protein modification provides a neat solution to this paradox: dynamic clusters naturally emerge during simulations, with constituent proteins recycling on a timescale proportional to the inverse switching rate, α^{-1} (Fig. 3). Importantly, when clusters in simulations are photobleached, they behave like nuclear bodies seen in vivo—they retain a memory of their shape, despite the continual exchange with the soluble pool.

Third and finally, switching affects large-scale chromatin organization. Bridging-induced clusters are associated with the formation of chromatin domains, reminiscent of the TADs observed in Hi-C data (19). Using a fiber patterned in such a way that it spontaneously folds into TADs, we find that switching has little effect on local TAD organization, but strongly suppresses inter-TAD interactions; local contacts are favored over nonlocal ones (Fig. 4). We expect that similar trends should be observed in more complex models for bridging-induced chromosome organization, such as those in Barbieri et al. (6) and Brackley et al. (8,20).

While here we focus on a flexible chromatin fiber, we expect that similar results should be found with a semiflexible one (20,22); then, our conclusions should also apply to bacterial DNA. Possible differences with respect to the results reported here may arise as a consequence of the first-order coil-globule phase transition displayed by stiffer chains (22). It would be interesting to study this case further in the future. Similar results to those reported here are also expected with more complex pathways between active and inactive states (e.g., modeling the cyclic flooding of proteins into nuclei, or their cyclic synthesis/degradation), and it would be of interest to investigate these scenarios.

We suggest that a mechanistic test of our model may be realized by disrupting some of the biochemical pathways involved in posttranslational protein modification, as this procedure would lead to a change in the switching rate α. For example, phosphorylation pathways (42,43) are known to mark protein degradation, hence decreasing the efficiency of phosphorylation may be expected to decrease α, thereby increasing typical cluster size and presumably affecting recovery time of FRAP experiments.

In summary, we demonstrated how nonequilibrium processes involving ephemeral protein states can provide a simple way of understanding how dynamic nuclear bodies of self-limiting size might form, and how chromosomal domains at the larger scale might be organized.

SUPPORTING MATERIAL

Supporting Materials and Methods, five figures, and six movies are available at http://www.biophysj.org/biophysj/supplemental/S0006-3495(17)30146-7.

AUTHOR CONTRIBUTIONS

C.A.B., B.L., D. Michieletto, F.M., P.R.C., and D. Marenduzzo designed the research, performed the research, analyzed the data, and wrote the article.

ACKNOWLEDGMENTS

This work was funded by the European Research Council (ERC) (under Consolidator Grant No. 648050 THREEDCELLPHYSICS), and by a Marie Skłodowska-Curie Intra European Fellowship (under grant No. 654908).

REFERENCES

1. Calladine, C. R., H. R. Drew, …, A. A. Travers. 2004. Understanding DNA, Ed. 3. Elsevier, London, UK.
2. Alberts, B. 2002. Molecular Biology of the Cell. Garland Science, New York.
3. Cook, P. R. 2001. Principles of Nuclear Structure and Function. Wiley-Liss, New York.
4. Dame, R. T., M. C. Noom, and G. J. Wuite. 2006. Bacterial chromatin organization by H-NS protein unravelled using dual DNA manipulation. *Nature.* 444:387–390.
5. Simonis, M., P. Klous, …, W. de Laat. 2006. Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). *Nat. Genet.* 38:1348–1354.
6. Barbieri, M., M. Chotalia, …, M. Nicodemi. 2012. Complexity of chromatin folding is captured by the strings and binders switch model. *Proc. Natl. Acad. Sci. USA.* 109:16173–16178.
7. Brackley, C. A., J. M. Brown, …, D. Marenduzzo. 2016. Predicting the three-dimensional folding of *cis*-regulatory regions in mammalian genomes using bioinformatic data and polymer models. *Genome Biol.* 17:59.
8. Brackley, C. A., J. Johnson, …, D. Marenduzzo. 2016. Simulated binding of transcription factors to active and inactive regions folds human chromosomes into loops, rosettes and topological domains. *Nucleic Acids Res.* 44:3503–3512.
9. Kilic, S., A. L. Bachmann, …, B. Fierz. 2015. Multivalency governs HP1α association dynamics with the silent chromatin state. *Nat. Commun.* 6:7313.
10. Harshman, S. W., N. L. Young, …, M. A. Freitas. 2013. H1 histones: current perspectives and challenges. *Nucleic Acids Res.* 41:9593–9609.
11. Wani, A. H., A. N. Boettiger, …, N. J. Francis. 2016. Chromatin topology is coupled to Polycomb group protein subnuclear organization. *Nat. Commun.* 7:10291.
12. Michieletto, D., D. Marenduzzo, and A. H. Wani. 2016. Chromosome-wide simulations uncover folding pathway and 3D organization of interphase chromosomes. arXiv:1604.03041.
13. Zhu, L., and C. P. Brangwynne. 2015. Nuclear bodies: the emerging biophysics of nucleoplasmic phases. *Curr. Opin. Cell Biol.* 34:23–30.
14. Sleeman, J. E., and L. Trinkle-Mulcahy. 2014. Nuclear bodies: new insights into assembly/dynamics and disease relevance. *Curr. Opin. Cell Biol.* 28:76–83.
15. Mao, Y. S., B. Zhang, and D. L. Spector. 2011. Biogenesis and function of nuclear bodies. *Trends Genet.* 27:295–306.
16. Pirrotta, V., and H. B. Li. 2012. A view of nuclear Polycomb bodies. *Curr. Opin. Genet. Dev.* 22:101–109.
17. Berry, J., S. C. Weber, …, C. P. Brangwynne. 2015. RNA transcription modulates phase transition-driven nuclear body assembly. *Proc. Natl. Acad. Sci. USA.* 112:E5237–E5245.
18. Brangwynne, C. P., P. Tompa, and R. V. Pappu. 2015. Polymer physics of intracellular phase transitions. *Nat. Phys.* 11:899–904.
19. Dixon, J. R., S. Selvaraj, …, B. Ren. 2012. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature.* 485:376–380.
20. Brackley, C. A., S. Taylor, …, D. Marenduzzo. 2013. Nonspecific bridging-induced attraction drives clustering of DNA-binding proteins and genome organization. *Proc. Natl. Acad. Sci. USA.* 110:E3605–E3611.
21. Johnson, J., C. A. Brackley, …, D. Marenduzzo. 2015. A simple model for DNA bridging proteins and bacterial or human genomes: bridging-induced attraction and genome compaction. *J. Phys. Condens. Matter.* 27:064119.
22. Le Treut, G., F. Képès, and H. Orland. 2016. Phase behavior of DNA in the presence of DNA-binding proteins. *Biophys. J.* 110:51–62.
23. Marenduzzo, D., and E. Orlandini. 2009. Topological and entropic repulsion in biopolymers. *J. Stat. Mech. Theory Exp.* 2009:L09002.
24. Cates, M. E., and T. A. Witten. 1986. Chain conformation and solubility of associating polymers. *Macromolecules.* 19:732–739.
25. Orlandini, E., and T. Garel. 1998. Collapse transitions of a periodic hydrophilic hydrophobic chain. *Eur. Phys. J. B.* 6:101–110.
26. Scolari, V. F., and M. Cosentino Lagomarsino. 2015. Combined collapse by bridging and self-adhesion in a prototypical polymer model inspired by the bacterial nucleoid. *Soft Matter.* 11:1677–1687.
27. Tootle, T. L., and I. Rebay. 2005. Post-translational modifications influence transcription factor activity: a view from the ETS superfamily. *BioEssays.* 27:285–298.
28. Nicodemi, M., and A. Prisco. 2009. Thermodynamic pathways to genome spatial organization in the cell nucleus. *Biophys. J.* 96:2168–2177.
29. Chaikin, P. M., and T. C. Lubensky. 2000. Principles of Condensed Matter Physics. Cambridge University Press, Cambridge, UK.
30. Byrne, A., P. Kiernan, …, K. A. Dawson. 1995. Kinetics of homopolymer collapse. *J. Chem. Phys.* 1012:573–577.
31. Cross, M. C., and P. C. Hohenberg. 1993. Pattern formation outside of equilibrium. *Rev. Mod. Phys.* 65:851–1112.
32. Zentner, G. E., and S. Henikoff. 2013. Regulation of nucleosome dynamics by histone modifications. *Nat. Struct. Mol. Biol.* 20:259–266.
33. Sheinman, M., O. Bénichou, …, R. Voituriez. 2012. Classes of fast and specific search mechanisms for proteins on DNA. *Rep. Prog. Phys.* 75:026601.
34. Mueller, F., D. Mazza, …, J. G. McNally. 2010. FRAP and kinetic modeling in the analysis of nuclear protein dynamics: what do we really know? *Curr. Opin. Cell Biol.* 22:403–411.

35. Papantonis, A., and P. R. Cook. 2013. Transcription factories: genome organization and gene regulation. *Chem. Rev.* 113:8683–8705.

36. Marenduzzo, D., C. Micheletti, and P. R. Cook. 2006. Entropy-driven genome organization. *Biophys. J.* 90:3712–3721.

37. Handwerger, K. E., C. Murphy, and J. G. Gall. 2003. Steady-state dynamics of Cajal body components in the *Xenopus* germinal vesicle. *J. Cell Biol.* 160:495–504.

38. Hernández-Muñoz, I., P. Taghavi, …, M. van Lohuizen. 2005. Association of BMI1 with polycomb bodies is dynamic and requires PRC2/EZH2 and the maintenance DNA methyltransferase DNMT1. *Mol. Cell. Biol.* 25:11047–11058.

39. Kimura, H., K. Sugaya, and P. R. Cook. 2002. The transcription cycle of RNA polymerase II in living cells. *J. Cell Biol.* 159:777–782.

40. Sanborn, A. L., S. S. Rao, …, E. L. Aiden. 2015. Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. *Proc. Natl. Acad. Sci. USA.* 112:E6456–E6465.

41. Sexton, T., E. Yaffe, …, G. Cavalli. 2012. Three-dimensional folding and functional organization principles of the *Drosophila* genome. *Cell.* 148:458–472.

42. Qian, H. 2007. Phosphorylation energy hypothesis: open chemical systems and their biological functions. *Annu. Rev. Phys. Chem.* 58:113–142.

43. Trentini, D. B., M. J. Suskiewicz, …, T. Clausen. 2016. Arginine phosphorylation marks proteins for degradation by a Clp protease. *Nature.* 539:48–53.

Biophysical Journal, Volume *112*

Supplemental Information

Ephemeral Protein Binding to DNA Shapes Stable Nuclear Bodies and Chromatin Domains

Chris A. Brackley, Benno Liebchen, Davide Michieletto, Francois Mouvet, Peter R. Cook, and Davide Marenduzzo

Ephemeral protein binding to DNA shapes stable nuclear bodies and chromatin domains: Supporting Information

C. A. Brackley, B. Liebchen, D. Michieletto, F. Mouvet, P. R. Cook, and D. Marenduzzo

Here we give more details on the simulations (including parameter values), and on the continuum mean field model (derivation, linear stability analysis and amplitude equation); we also show additional results and figures which are discussed in the main text.

I. DETAILS OF BROWNIAN DYNAMICS SIMULATIONS

The chromatin fiber is modeled as a bead-spring polymer with finitely-extensible non-linear elastic springs via a Kremer-Grest model [1]. To map length scales from simulation to physical units, we can, e.g., set the diameter, σ, of each bead to ~ 30nm$\simeq 3$ kbp (assuming an underlying 30 nm fiber; of course, all our results would remain valid with a different mapping).

Letting $\boldsymbol{r}_i$ and $\boldsymbol{d_{i,j}} \equiv \boldsymbol{r}_j - \boldsymbol{r}_i$ be respectively the position of the centre of the i-th bead and the vector of length $d_{i,j}$ between beads i and j, we can express the potential modeling the connectivity of the chain as

$$U_{\mathrm{FENE}}(i,i+1) = -\frac{k}{2}R_0^2 \ln\left[1-\left(\frac{d_{i,i+1}}{R_0}\right)^2\right],$$

for $d_{i,i+1} < R_0$ and $U_{\mathrm{FENE}}(i,i+1) = \infty$, otherwise; here we chose $R_0 = 1.6\ \sigma$ and $k = 30\ k_BT/\sigma^2$.

The bending rigidity of the chain is described through a standard Kratky-Porod potential, as follows

$$U_b(i,i+1,i+2) = \frac{k_BTl_p}{\sigma}\left[1-\frac{\boldsymbol{d}_{i,i+1}\cdot\boldsymbol{d}_{i+1,i+2}}{d_{i,i+1}d_{i+1,i+2}}\right],$$

where we set the persistence length $l_p = 3\sigma \simeq 90$ nm, which is reasonable for a chromatin fiber.

The steric interaction between a chromatin bead, a, and a protein bridge, b (of sizes $\sigma_a = \sigma_b = \sigma$), is modeled through a truncated and shifted Lennard-Jones potential

$$U_{\mathrm{LJ}}(i,j) = 4\epsilon_{ab}\left[\left(\frac{\sigma}{d_{i,j}}\right)^{12} - \left(\frac{\sigma}{d_{i,j}}\right)^{6} - \left(\frac{\sigma}{r_c}\right)^{12} + \left(\frac{\sigma}{r_c}\right)^{6}\right],$$

for $d_{i,j} < r_c$ and 0 otherwise. This parameter, r_c, is the interaction cutoff; it is set to $r_c = 2^{1/6}\sigma$ for inactive proteins, in order to model purely repulsive interactions, and to $r_c = 1.8\sigma$ for an active protein, so as to include attractive interactions. In both cases, the potential is shifted to zero at the cut-off in order to have a smooth curve and avoid singularities in the forces. Purely repulsive interactions, such as those between inactive proteins and chromatin segments, are modeled by setting $\epsilon_{ab} = k_BT$, while attractive interactions are modeled using: (i) $\epsilon_{ab} = 3k_BT$ (for non-specific interactions, Fig. 1); (ii) $\epsilon_{ab} = 15k_BT$ and $\epsilon_{ab} = 4k_BT$ (for non-specific and specific interactions respectively, Fig. 3); (iii) $\epsilon = 4k_BT$ (for non-specific interactions, Fig. 4); or (iv) as specified in Supporting Figure captions in other cases.

The total potential energy experienced by bead i is given by

$$\begin{aligned} U_i \;=\; & \sum_j U_{\mathrm{FENE}}(i,j)\delta_{j,i+1} + \\ & \sum_j\sum_k U_{\mathrm{b}}(i,j,k)\delta_{j,i+1}\delta_{k,i+2} + \sum_j U_{\mathrm{LJ}}(i,j), \end{aligned} \tag{1}$$

and its dynamics can be described by the Langevin equation

$$m\ddot{\boldsymbol{r}}_i = -\xi\dot{\boldsymbol{r}}_i - \nabla U_i + \boldsymbol{\eta}_i, \tag{2}$$

where m is the bead mass, ξ is the friction coefficient, and $\boldsymbol{\eta}_i$ is a stochastic delta-correlated noise. The variance of each Cartesian component of the noise, σ_η^2, satisfies the usual fluctuation dissipation relation $\sigma_\eta^2 = 2\xi k_BT$.

As is customary [1], we set $m/\xi = \tau_{\rm LJ} = \tau_{\rm B}$, with the LJ time $\tau_{\rm LJ} = \sigma\sqrt{m/\epsilon}$ and the Brownian time $\tau_{\rm B} = \sigma/D_b$, where ϵ is the simulation energy unit, equal to k_BT, and $D_b = k_BT/\xi$ is the diffusion coefficient of a bead of size σ. From the Stokes friction coefficient for spherical beads of diameter σ we have that $\xi = 3\pi\eta_{sol}\sigma$ where η_{sol} is the solution viscosity. One can map this to physical units by setting the viscosity to that of the nucleoplasm, which ranges between $10 - 100$ cP, and by setting $T = 300$ K and $\sigma = 30$ nm, as above. From this it follows that $\tau_{\rm LJ} = \tau_{\rm B} = 3\pi\eta_{sol}\sigma^3/\epsilon \simeq 0.6 - 6$ ms; τ_B is our time simulation unit, used when measuring time in the figures in the main text and in this Supporting Information. The numerical integration of Eq. (2) is performed using a standard velocity-Verlet algorithm with time step $\Delta t = 0.01\tau_{\rm B}$ and is implemented in the LAMMPS engine. We perform simulations for up to 2×10^5 $\tau_{\rm B}$, which correspond to 2-20 minutes in real time. Protein switching is including by coupling an external code to LAMMPS; the external code changes stochastically type with rate α. This code is called every 1000 or 10000 LAMMPS Brownian dynamics steps, through the LAMMPS input file.

II. MEAN FIELD THEORY FOR SWITCHING PROTEINS

In our particle based simulations we observed the growth of clusters due to bridging interactions (see main text). When protein activation-inactivation reactions were absent, these clusters coarsened, resulting in one large macroscopic cluster in steady state. However, in the presence of these reactions, the clusters coarsened only up to a self-limiting size. To better understand this transition from macrophase separation to microphase separation, and the involved length scales, we now develop a phenomenological minimal model for the dynamics of chromatin and proteins. We describe the distribution of chromatin via the probability density field $\rho(\mathbf{x}, t)$, and the density of active, or binding, and inactive, or non-binding, proteins by $\Phi_a(\mathbf{x}, t) \equiv \Phi(\mathbf{x}, t)$ and $\Phi_i(\mathbf{x}, t)$ respectively.

The starting point for our model is the free energy $\mathcal{F} = \int f(\mathbf{x})\mathrm{d}\mathbf{x}$ where f is the free energy density:

$$f = \frac{D_1'}{2}\rho^2 + \frac{D_2'}{2}\Phi^2 - \chi'\rho\Phi + \frac{k'}{2}(\nabla\rho)^2 + \frac{g'}{4}\rho^4. \tag{3}$$

Here, the first two terms describe diffusion of chromatin and proteins respectively, the third term describes the energy gain through bridging and the last two terms, multiplied by k', g', respectively penalize sharp interfaces due to interfacial tension, and strong accumulations of chromatin due to short ranged repulsions.

Assuming diffusive dynamics here and using the fact that in the absence of protein modification, the number density of all species (ρ, Φ, Φ_i) is conserved, we can derive the equations of motions for our fields as done for model B dynamics [2]. However, in the presence of active protein modification, we need an additional reaction term, so that our equations of motion read

$$\dot\rho = M_\rho\nabla^2\frac{\delta\mathcal{F}}{\delta\rho}, \tag{4}$$

$$\dot\Phi_a = M_a\nabla^2\frac{\delta\mathcal{F}}{\delta\Phi_a} - \alpha\Phi_a + \beta\Phi_i. \tag{5}$$

Here M_ρ and M_a are dimensionless mobility coefficients of chromatin and activated proteins respectively, while α and β are the activation and inactivation rates for proteins. Since inactive proteins do not bind, we assume that they diffuse quickly, i.e. that their density field is uniform.

Now integrating Eq. (5) over the whole system and denoting the total number of active and inactive proteins with $N_a(t)$ and $N_i(t)$ respectively, we obtain $\dot N_a = -\alpha N_a + \beta N_i$. Conservation of the total protein number $N = N_a + N_b$ now yields $\dot N_i = (1 + \beta/\alpha)N_i$ which approaches the steady state $N_i = \alpha N/(\alpha + \beta)$, i.e. $\Phi_i = \alpha/(\alpha + \beta)$, exponentially fast. Now defining $\Phi_0 := (\beta/\alpha)\Phi_i = \beta/(\alpha + \beta)$ (and ignoring short-time effects due to possible 'imbalances' between active and inactive proteins in the initial state), Eqs. (4,5) reduce to:

$$\dot\rho = M_\rho\nabla^2[a_1\rho - k\nabla^2\rho - \chi\Phi + g\rho^3], \tag{6}$$

$$\dot\Phi = M_\Phi\nabla^2[a_2\Phi - \chi\rho] - \alpha(\Phi - \Phi_0), \tag{7}$$

where for simplicity hereon we drop the subscript a on Φ_a for active proteins. We also introduced $D_1 = M_\rho a_1$ and $D_2 = M_\Phi a_2$.

To further reduce these equations and to identify a minimal set of dimensionless control parameters, we now choose time and space units $t_u = 1/\alpha$ and $x_u = \sqrt{D_2/\alpha}$ and redefine $\Phi = \Phi\chi M_\rho/D_2$. This leads to

$$\dot\rho = \mathcal{D}_0\nabla^2\rho - A\nabla^4\rho - \nabla^2\phi + G\nabla^2\rho^3, \tag{8}$$

$$\dot\Phi = \nabla^2\Phi - X\nabla^2\rho - (\Phi - \Phi_0). \tag{9}$$

That is, our parameter space is spanned by the four dimensionless numbers $X = (\chi^2 M_\rho M_\Phi / D_2^2)$; $\mathcal{D}_0 = (D_1/D_2)$; $A = \alpha k M_\rho/(D_2^2)$ and $G = g M_\rho / D_2$.

A. Linear stability analysis

To better understand in which parameter regimes we should expect (i) a uniform distribution of chromatin and proteins, (ii) cluster growth proceeding to macroscopic phase separation and (iii) microphase separation, we now perform a linear stability analysis. This analysis will equip us with a prediction for the self-limiting cluster size in regime (iii), matching the results of our particle based simulations. We therefore study the response of the uniform phase to small perturbations in the density fields (ρ, Φ). Linearising Eqs. (8,9) around the uniform solution $(\rho, \Phi) = (\rho_0, \Phi_0)$, where ρ_0 is the DNA density as fixed by the initial state, leads to the following equations of motion for the fluctuations $\rho' = \rho - \rho_0, \Phi' = \Phi - \Phi_0$:

$$\dot{\rho}' = \mathcal{D}\nabla^2\rho' - A\nabla^4\rho - \nabla^2\Phi', \tag{10}$$

$$\dot{\Phi}' = \nabla^2\Phi' - X\nabla^2\rho' - \Phi'. \tag{11}$$

Here, we defined $\mathcal{D} := \mathcal{D}_0 + 3G\rho_0^2$. Fourier transforming Eqs. (10,11) and using $Q := \mathbf{q}^2$ leads to the following dispersion relation (or characteristic polynomial),

$$\lambda(Q) = \frac{1}{2}\Big[-1 - Q\,(1 + \mathcal{D} + AQ) \pm\sqrt{[-1 + Q(\mathcal{D} - 1 + AQ)]^2 + 4Q^2X}\Big], \tag{12}$$

which links the growth rate λ of the fluctuation with its wavevector Q. An analysis of this relation leads us to the instability criterion

$$\sqrt{X} > \sqrt{X_C} := \sqrt{A} + \sqrt{\mathcal{D}}, \tag{13}$$

which translates, in physical units, to

$$\chi > \sqrt{\frac{k\alpha}{M_\Phi}} + \sqrt{\frac{D_2}{M_\Phi}\left[\frac{D_1}{M_\rho} + 3g\rho_0^2\right]}. \tag{14}$$

This criterion determines the transition line (hypersurface) between regions (i) and (ii/iii) in the parameter space. Hence, if the bridging interactions are sufficiently large, small fluctuations around the uniform state will grow to form clusters. Remarkably, this instability and the corresponding emergence of order (clustered phase) is not contingent on the presence of a certain minimal protein (or DNA) density, suggesting that even a very low protein concentration is sufficient to trigger clustering.

To map out the transition line from macrophase separation to microseparation (at the onset of instability), it is useful to consider the wavelength at which instability first occurs. From Eq. (12) and $q_c = \partial_q \lambda(q) = 0|_{X=(\sqrt{A}+\sqrt{D})^2}$, we find $q_c = (D/A)^{1/4}$, corresponding, in physical units, to the length scale

$$L_c = \frac{2\pi}{q_c} = 2\pi\left(\frac{D_1 D_2 + 3M_\rho g\rho_0^2 D_2}{\alpha k M_\rho}\right)^{1/4}. \tag{15}$$

Thus, in an infinite system, coarsening only occurs for $\alpha = 0$. [In finite systems macrophase separation is observed if α is small enough that L_c exceeds the system size.] From this analysis we expect the average particle number per cluster to scale as $N \propto L_c^3 \propto \alpha^{-3/4}$ (at least close to the onset of instability). This value agrees well with the numerically observed scaling of $N \propto \alpha^{-0.76}$ (Fig. 2B), supporting the view that the essential physics of chromatin clustering can be described and understood within our simplified mean field theory.

For completeness, we also calculate the boundaries of the instability band from Eq. (12), which, after translating back into physical units (for $M_1 = M_2 = 1$), read as follows:

$$q_\pm = \frac{1}{\sqrt{2D_2K}}\sqrt{\nu \pm \sqrt{\nu^2 - 4D_2K\alpha(D_1 + 3g\rho_0^2)}}, \tag{16}$$

$$\nu = \left[\chi^2 - D_1D_2 - 3D_2g\rho_0^2 - K\alpha\right]. \tag{17}$$

At the onset of instability, we find $\nu \to 4D_2K\alpha(D_1 + 3g\rho_0)$ and hence we recover the $\alpha^{1/4}$-scaling of the onset mode. In contrast, the boundaries of the instability band scale in a more complicated way which is nonuniversal in α.

B. Amplitude equations

We now perform a perturbative analysis of the linearly unstable modes (fluctuations) close to onset of instability. This analysis will lead us to a further reduced effective model, describing the linear growth and nonlinear saturation of chromatin clusters on large scales and at long timescales.

We begin by rewriting Eqs. (4,5) as

$$\mathcal{L}\begin{pmatrix}\rho'\\ \Phi'\end{pmatrix}+\mathcal{N}-\begin{pmatrix}\dot\rho'\\ \dot\Phi'\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix}, \tag{18}$$

where the linear operator $\mathcal{L}$ and the nonlinear term $\mathcal{N}$ represent

$$\mathcal{L}=\begin{pmatrix}\mathcal{D}\partial_x^2-A\partial_x^4 & -\partial_x^2\\ -X\partial_x^2 & \partial_x^2-1\end{pmatrix};\quad \mathcal{N}=G\begin{pmatrix}\partial_x^2\rho'^3+3\rho_0\partial_x^2\rho'^2\\ 0\end{pmatrix}. \tag{19}$$

Now, we replace (as usual, see [3])

$$X\to(1+\epsilon)X_C;\quad \partial_x\to\partial_x+\sqrt{\epsilon}\partial_X;\quad \partial_t\to\epsilon\partial_T \tag{20}$$

where $\epsilon=(X-X_C)/X_C$ and expand the fields as

$$\rho'=\sum_{n=1}^{\infty}\epsilon^{n/2}\rho_{n-1};\quad \Phi'=\sum_{n=1}^{\infty}\epsilon^{n/2}\Phi_{n-1}. \tag{21}$$

Next, we plug these expansions into Eqs. (19) and solve the resulting equations to lowest order ($\epsilon^{1/2}$). Using the Ansatz $\rho_0=\mathcal{A}\exp(\mathrm{i}q_cx)+c.c.$ and $\Phi_0=\mathcal{A}_\phi\exp(\mathrm{i}q_cx)+c.c.$ with amplitudes $\mathcal{A},\mathcal{A}_\phi$, we find $q_c=(\mathcal{D}/A)^{1/4}$ reproducing the corresponding result from our linear stability analysis (see above), as well as $\mathcal{A}_\phi=\mathcal{A}(\mathcal{D}+Aq_c^2)=\mathcal{A}\sqrt{\mathcal{D}X_C}$ which fixes the relation between the amplitudes of both density fields. The solution of our perturbative equations to order $\epsilon^{1/2}$ then reads $\rho'=2\mathcal{A}\cos q_cx$ with the so-far unknown amplitude $\mathcal{A}$.

The result to order ϵ turns out not to be particular useful for our purpose, as solving it would provide us with a similar result as to order $\epsilon^{1/2}$), but with another unknown amplitude $\mathcal{A}'$ yielding a higher order correction to the solution $\rho'=2\mathcal{A}\cos q_cx$. Since we are looking only for the lowest order result in ϵ we directly consider the perturbative equations of motion to order $\epsilon^{3/2}$. As usual [3], we do not attempt to solve the corresponding equations explicitly, but apply Fredholm's theorem providing solvability conditions, which determine an equation of motion for $\mathcal{A}$. After a long but straightforward calculation and transforming back to coordinates t,x we find:

$$c_t\dot{\mathcal{A}}=\epsilon\mathcal{A}+c_x\partial_x^2\mathcal{A}+c_3\mathcal{A}^3, \tag{22}$$

where

$$c_t = \sqrt{\frac{A}{X_C}}\left(1+\frac{1}{\mathcal{D}}\right), \tag{23}$$

$$c_x = \frac{4A}{\sqrt{\mathcal{D}X_C}}, \tag{24}$$

$$c_3 = \frac{3G}{\sqrt{\mathcal{D}X_C}}. \tag{25}$$

Eq. (22) is a variant of the real Ginzburg-Landau equation, here describing, together with the coefficients Eqs. (23–25), the dynamics of chromatin and proteins close to the onset of instability. In this equation $\epsilon/(c_tt_u)$ is the initial growth rate of protein clusters; $x_u\sqrt{\epsilon/c_3}$ describes the amplitude of their saturation (related to their density) for a given $X>\left[\sqrt{A}+\sqrt{\mathcal{D}}\right]^2$ and $x_u\sqrt{c_x}$ is a correlation length, describing a scale of spatial modulations of the saturation amplitude of DNA clusters.

Although we equipped our original equilibrium model with reaction terms which drive it out of equilibrium, its large scale and long time dynamics (i.e., Eq. (22)) can be effectively mapped (at least close to onset of instability) onto a potential system with the following Lyapunov functional:

$$\mathcal{V}[\mathcal{A}]=\int\mathrm{d}x\left[-\epsilon|\mathcal{A}|^2+\frac{c_3}{2}|\mathcal{A}|^4+c_x^2|\partial_X\mathcal{A}|^2\right], \tag{26}$$

$$\dot{\mathcal{A}} = -\frac{1}{c_T}\frac{\delta V}{\delta \mathcal{A}}. \tag{27}$$

Hence, quite remarkably, the dynamics of the present reaction-diffusion system can be mapped, within this approximation, onto a system which is purely relaxational.

III. ADDITIONAL SIMULATION RESULTS

In this section we present additional simulation results, which complement those discussed in the main text.

Figure S1 shows that the FRAP signal (following simulated photobleaching of a spherical spot of size 50 σ) shows recovery also for equilibrium bridges, if the specific and non-specific interactions are carefully tuned. However, protein modification provides a more robust way to achieve this, which simultaneously allows stable binding (when the protein is in the active state), and fast turnover (due to the unbinding and diffusion of inactive proteins).

Figure S2 shows the cluster size for different parameter values for the case of non-specific protein-chromatin interactions. This demonstrates that it can be varied significantly (by about an order of magnitude), and is particularly sensitive to the protein concentration.

Figures S3 and S4 highlight some further properties of the recycling clusters. In particular, Figure S3 shows that these clusters retain memory of their shape even as the proteins which constitute them change. Figure S4 shows the dynamics of some protein and chromatin beads with and without modification. Without modification, once proteins bind to a cluster they diffuse little for the rest of the simulation, whereas with modification they sample the whole simulation domain. Contrary to this, the dynamics of the chromatin beads within a cluster is similar with and without modification: they diffuse very little. This explains why clusters keep their shape: while proteins bind and unbind, the underlying chromatin backbone is largely unchanged.

Finally, Figure S5 shows how the effect of protein switching on the ratio between non-local and local contacts, shown in Figure 4 in the main text, is affected by the values of non-specific and specific interactions.

IV. ESTIMATES OF RELEVANT TIMESCALES

Here we provide a series of simple estimates for the value of the relevant timescales in our problems. Consider first a fluorescence-recovery-after-photobleaching, or FRAP, experiment, where a cluster of size $\sigma_{\rm cl} \sim 0.1 - 1$ μm is inside the bleached spot, which we imagine has a diameter of $\sigma_{\rm FRAP} \sim 1$ μm. In this Section, as previously, σ will instead denote the size of a typical chromatin-binding complex, or chromatin bead (as previously, we imagine this is ~ 30 nm).

What is the timescale for the recovery of the FRAP signal? Clearly, this depends on the underlying dynamics of the bleached/unbleached proteins. If proteins diffuse freely, then unbleached proteins can enter the bleach spot to give recovery within a time, $\tau_{\rm diff}$, proportional to

$$\tau_{\rm diff} \sim \frac{\sigma^2_{\rm FRAP}}{D}. \tag{28}$$

For a protein size $\sigma \sim 30$ nm, and if the nucleoplasm viscosity is 10 cP (ten times that of water), the diffusion coefficient is $\sim 1.4\mu{\rm m}^2$ s-1, so that $\tau_{\rm diff} \sim 1$s, which is too fast to account for FRAP response of nuclear bodies (furthermore, of course, freely diffusing proteins could not self-organise into clusters).

If, instead, non-switching binding proteins create a cluster, then the FRAP signal recovers when some proteins unbind, and others replace these from the soluble (unbleached) pool. As the former process is slower than the latter (which relies again on diffusion), we can equate the FRAP recovery timescale to the time needed for an equilibrium protein to unbind from the cluster, which can be estimated as,

$$\tau_{\rm non-switch} \sim \frac{\sigma^2_{\rm cl}}{D}\exp\left(\frac{\Delta U}{k_B T}\right) \tag{29}$$

where ΔU indicates the strength of chromatin-protein interaction. If we assume an interaction of 10 kcal/mol, consistent with either multiple non-specific or a single specific DNA-protein interactions, then $\tau_{\rm non-switch} > 10^5$ s, which is too slow to account for the FRAP recovery observed in nuclear bodies. Clearly, changing ΔU will change $\tau_{\rm non-switch}$, but in order for the estimate to be in the observed range, the interaction energy would have to be finely tuned, and would be significantly lower than that seen in typical DNA-protein interactions.

If, finally, switching proteins are in the cluster, then the unbinding time, which again can be equated to the FRAP recovery time, is simply

$$\tau_{\rm switch} \sim \alpha^{-1}. \tag{30}$$

For typical post-translational modification, or transcription termination, this is in the several seconds to minutes timescale, which is compatible with experimental results.

Aside from FRAP, another important timescale is that over which local TADs form (e.g., in Fig. 4), $\tau_{\rm TAD}$. In analogy with polymer collapse and heteropolymer folding (see, e.g., Ref. [4]), we expect $\tau_{\rm TAD}$, to be a power law in the number of monomers in the TAD, say M, where the prefactor should describe microscopic (diffusion) dynamics of a monomer. Dimensional analysis then suggests

$$\tau_{\rm TAD} \sim \frac{\sigma^2}{D} M^z \sim \tau_{\rm B} M^z \tag{31}$$

where z is a scaling exponent. The Brownian time $\tau_{\rm B}$ is of the order of 10^{-3} s with previous assumptions for viscosity and monomer size, while in our simulations $z \simeq 1$ at least up to $M \sim 100$ (corresponding to 300 kbp). Also for eukaryotic chromosomes, TAD size is between 100 kbp and 1 Mbp, so M is at most a few hundred. Therefore, if $z = 1$, we estimate $\tau_{\rm TAD}$ to be of the order of 1 s, smaller than typical modification times – even assuming a larger effective value of z (e.g., $z = 2$ gives at most $\tau_{\rm TAD}$ of order of 1 min). Previous large-scale simulations also confirm that eukaryotic TADs form in minutes [5, 6]. These estimates explain why switching proteins in our simulations can still form TADs in pretty much the same way as non-switching proteins, and suggest that the same should also hold for real chromosomes.

V. CAPTIONS OF SUPPLEMENTARY MOVIES

Supplementary Movie 1: A movie of the simulation shown in Figure 1C of the main text. Proteins do not switch ($\alpha = 0$). First a snapshot 10^4 simulation units after equilibration is shown: a number of small clusters have formed. Then the subsequent dynamics are shown: clusters grow and merge, and coarsening proceeds indefinitely.

Supplementary Movie 2: A movie of the simulution shown in Figure 1D of the main text. Proteins switch at a rate $\alpha = 0.0001$ inverse Brownian times. Switching arrests coarsening, and leads to clusters of self-limiting size in steady state.

Supplementary Movie 3: Parameters for this Movie are as in Figure 3 of the main text for the $\alpha = 0$ case. Chromatin beads are not shown for simplicity. The movie starts with clusters which have formed during 10^4 simulation units following equilibration. The proteins are colored according to the cluster they belong to when the movie starts; proteins not in any clusters at that time are gray. The movie then follows the dynamics with non-switching proteins, for another 10^4 simulation units: it can be seen that colored clusters persist, therefore photobleaching such a cluster would lead to little or no recovery of signal in the cluster.

Supplementary Movie 4: As Supplementary Movie 3, but now with switching proteins ($\alpha = 0.0001$ inverse Brownian times). Proteins are colored according to the initial clusters; by the end of the simulations all clusters have mixed colors. While proteins in clusters recycle, the cluster retains the same overall shape.

Supplementary Movie 5: As in Supplementary Movie 4, but a zoom on two clusters to show more clearly clusters retain a "memory" of their shape.

Supplementary Movie 6: A movie of the simulation shown in Figure 4 in the main text. The first half of the simulation involves non-switching proteins and lasts 10^5 simulation units: two clusters form. Proteins are black; yellow chromatin beads are binding, while blue ones are non-binding. During the second half, proteins are able to switch ($\alpha = 0.0001$ inverse Brownian times); clusters split and interdomain interactions are suppressed.

[1] Kremer K., Grest G. S. (1990) Dynamics of entangled linear polymer melts: A molecular-dynamics simulation. *J. Chem. Phys.* 92(8):5057.

[2] P.M. Chaikin, P. M., Lubensky, T. C. (2000). *Principles of condensed matter physics*, Cambrdige University Press, Cambridge.
[3] Cross, M. C., Hohenberg, P. C. Pattern formation outside of equilibrium. *Rev. Mod. Phys* **65**, 851-1112 (1993).
[4] Byrne, A., Kiernan, P., Green, D., Dawson, K. A. (1995). Kinetics of homopolymer collapse. *J. Chem. Phys.* **1012**, 573-577 (1995).
[5] Michieletto, D., Marenduzzo, D., Wani, A. H. (2016) Chromosome-wide simulations uncover folding pathway and 3D organization of interphase chromosomes. arXiv:1604.03041.
[6] Brackley, C. A., Johnson, J., Kelly, S., Cook, P. R., Marenduzzo, D. (2016). Simulated binding of transcription factors to active and inactive regions folds human chromosomes into loops, rosettes and topological domains. *Nucl. Acids Res.* **44**, 3503-3512.

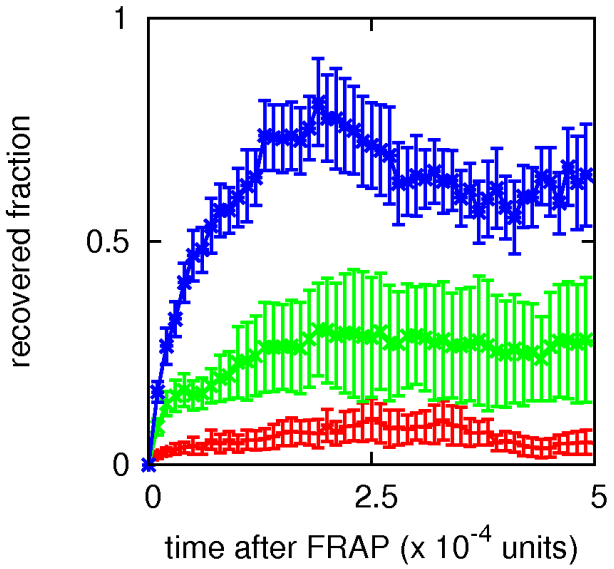

FIG. S1: **Comparison of FRAP recovery for non-switching and switching proteins**. FRAP recovery, measured as the number of unbleached proteins which are in the bleached volume after bleaching. The signals are normalized with the number of proteins in the bleached volume at the time of bleaching. As in Fig. 3Bi, the bleached volume is a sphere of size 50σ. Error bars give SD of mean, and time is given in multiples of 10^4 simulation units. Values of the specific and non-specific interactions, and of α, were respectively: $15k_BT$, $4k_BT$, 0 (red curve), $8k_BT$, $3k_BT$, 0 (green curve), and $15k_BT$, $4k_BT$, 0.0001 inverse Brownian times (blue curve). It can be seen that varying the values of non-specific and specific interactions can lead to FRAP recovery also for $\alpha = 0$ (green curve), although, in the absence of fine tuning, this is to a smaller extent with respect to $\alpha \neq 0$ (blue curve).

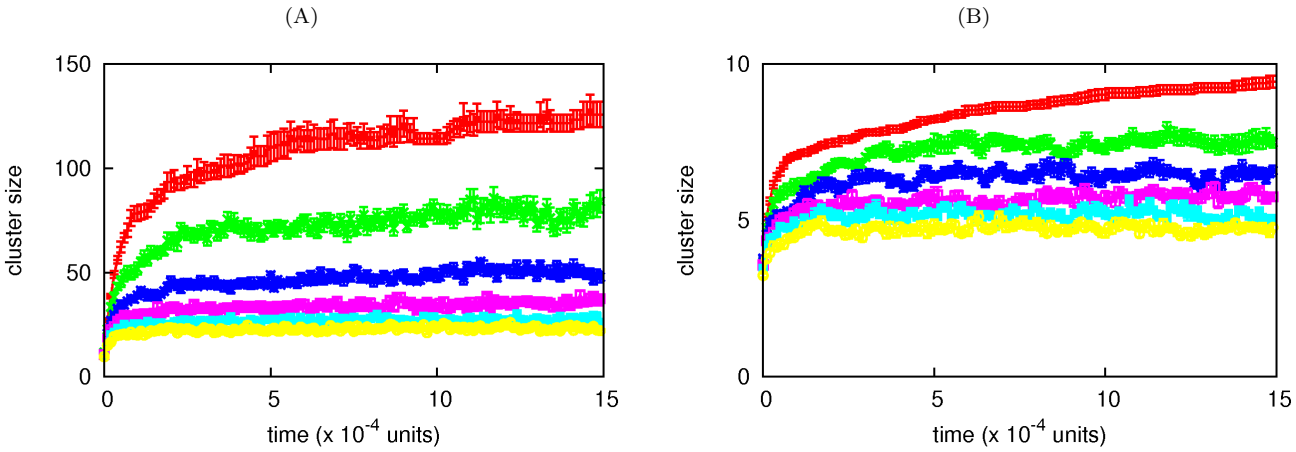

FIG. S2: **Cluster size with specific binding.** (A) Plot of the average number of proteins in a cluster versus time (± SD), for $N = 2000$ switching proteins binding to the chromatin fiber, both specifically (interaction strength 15 k_BT, cut-off 1.8σ), to every 20-th bead in the polymer, and non-specifically (interaction strength 4 k_BT, cut-off 1.8σ) to any other bead. From top to bottom, curves correspond to $\alpha = 0$ (in which case half of the proteins are non-binding, and half binding), 0.0001, 0.0002, 0.0003, 0.0004, 0.0005 respectively. (B) Same as (A), but now for $N = 500$ switching proteins, with specific interaction strength of $8k_BT$ and non-specific interaction of $3k_BT$; the interaction cut-off is 1.8σ.

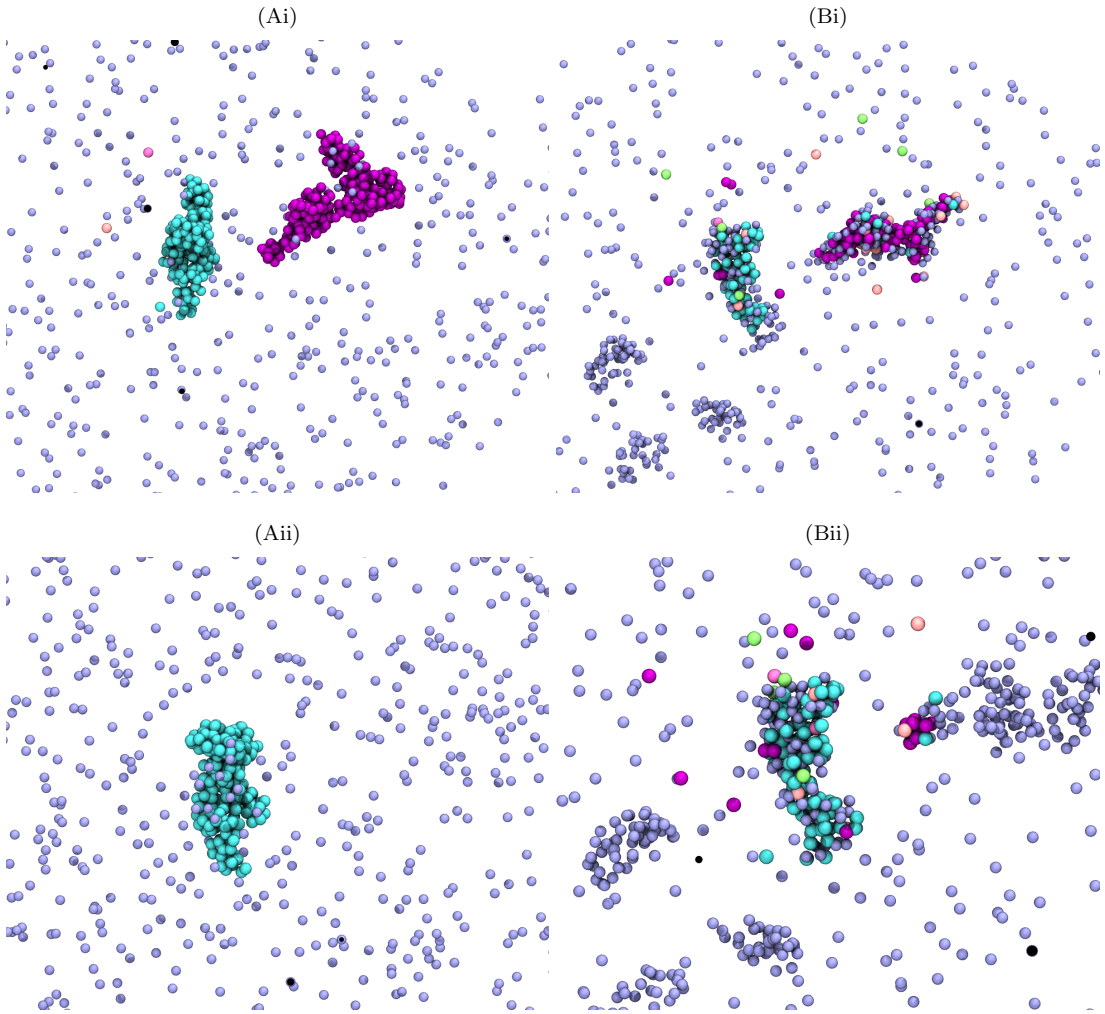

FIG. S3: **Switching proteins form clusters which retain memory of their shape.** This figure follows the evolution of clusters in a simulation analogous to that of Fig. 3A in the main text; the same parameters apply. Only proteins – and not chromatin beads – are shown for clarity. (A) Snapshots taken 10^4 time units after equilibration, for non-switching proteins, showing two clusters (beads are colored according to the cluster they belong to); (ii) shows another cluster. (B) Snapshots of the same regions shown in (A) after another 10^5 simulation units, and after allowing the proteins to now switch ($\alpha = 0.0001\tau_{\mathrm{B}}^{-1}$). Clusters recycle their constituent proteins whilst retaining a very similar shape.

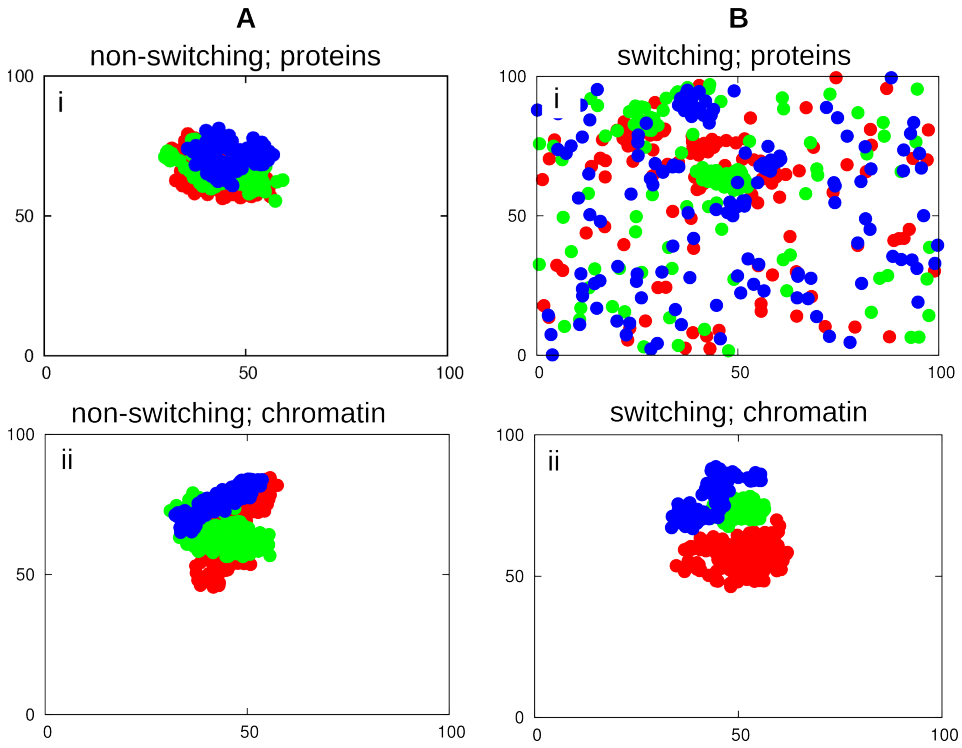

FIG. S4: **Trajectories of proteins and high-affinity chromatin beads.** Simulations are as in Fig. 3 of the main text; the same parameters apply. Positions of proteins and chromatin beads are shown in a 2D projection of the simulation domain, positions on the axes are measured in units of σ. (A) Non-switching proteins. (i) Red, green and blue circles denote positions of three non-switching proteins, recorded every 100 τ_B in a simulation (total length 1.5×10^5 simulation units. In this case, all three proteins remain bound to one cluster throughout the time series. (ii) Red, green and blue circles denote positions of three high affinity chromatin beads, again recorded every 100 τ_B in the same simulation. All three chromatin beads remain in the same cluster. (B) Same as (A), but for switching proteins ($\alpha = 0.0001$ inverse Brownian times). Now the three switching proteins diffuse through the whole space, while the three chromatin beads are still confined; this shows that the underlying scaffold of the cluster persists as the proteins are recycled.

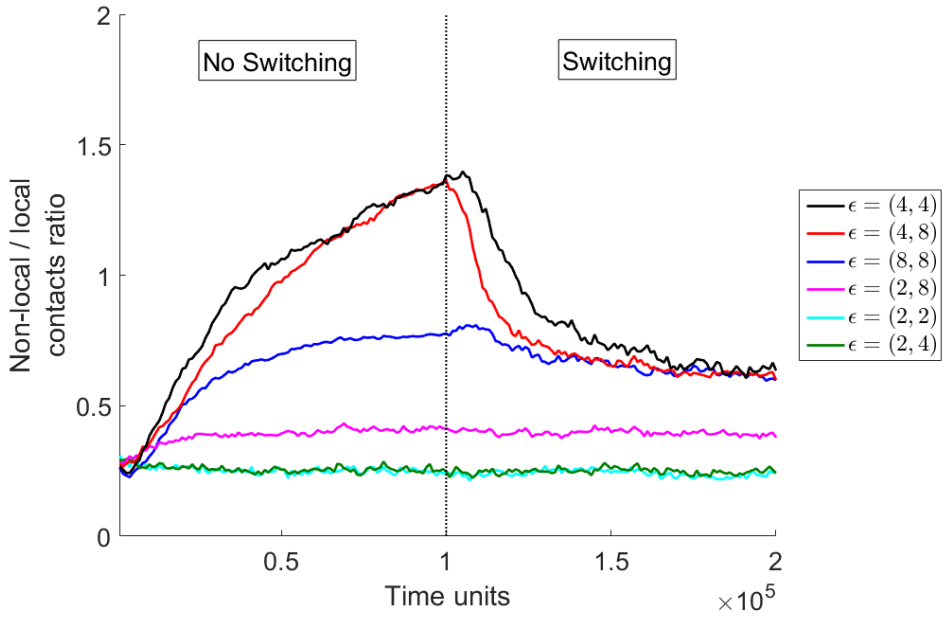

FIG. S5: **Protein switching favours local over non-local chromatin contacts.** The plot shows the fraction of non-local versus local contacts for a chromatin fiber; fiber patterning and all parameters are as in Fig. 4 of the main text. Simulations initially involved non-switching proteins; half-way through the simulation, proteins began to switch ($\alpha = 0.0001$ inverse Brownian times). Contacts are classified as local (non-local) if they involve beads separated less than (more than) 400 beads along the chain (or 1.2 Mbp. Non-specific (ϵ_1) and specific (ϵ_2) interaction energies are indicated on the right of the plot, in the format (ϵ_1, ϵ_2).