Epigenetic Transitions and Knotted Solitons in Stretched Chromatin

D. Michieletto[1,*], E. Orlandini[2], D. Marenduzzo[1]
[1] *SUPA, School of Physics and Astronomy, University of Edinburgh, Peter Guthrie Tait Road, Edinburgh, EH9 3FD, UK.* [2] *Dipartimento di Fisica e Astronomia and Sezione INFN, Universitá di Padova, Via Marzolo 8, Padova 35131, Italy*
corresponding author: davide.michieletto@ed.ac.uk

The spreading and regulation of epigenetic marks on chromosomes is crucial to establish and maintain cellular identity. Nonetheless, the dynamical mechanism leading to the establishment and maintenance of a given, cell-line specific, epigenetic pattern is still poorly understood. In this work we propose, and investigate *in silico*, a possible experimental strategy to illuminate the interplay between 3D chromatin structure and epigenetic dynamics. We consider a set-up where a reconstituted chromatin fibre is stretched at its two ends (e.g., by laser tweezers), while epigenetic enzymes (writers) and chromatin-binding proteins (readers) are flooded into the system. We show that, by tuning the stretching force and the binding affinity of the readers for chromatin, the fibre undergoes a sharp transition between a stretched, epigenetically disordered, state and a crumpled, epigenetically coherent, one. We further investigate the case in which a knot is tied along the chromatin fibre, and find that the knotted segment enhances local epigenetic order, giving rise to "epigenetic solitons" which travel and diffuse along chromatin. Our results point to an intriguing coupling between 3D chromatin topology and epigenetic dynamics, which may be investigated via single molecule experiments.

Introduction

Cell-line-specific features in multi-cellular organisms are achieved by regulation of "epigenetic marks", biochemical modifications of DNA and histone octamers which do not affect the underlying genomic sequence. Thus, the average pattern of epigenetic marks in a given cell line well correlates with the pattern of genes which are transcriptionally active or inactive within that cell line [1–3]. Dissecting the biophysical mechanisms leading to the *de novo* establishment, spreading and maintainance of epigenetic marks is consequently a key step towards a better understanding of the dynamic organisation of genomes and of chromosomal re-arrangement throughout the cell cycle [4], in cellular ageing [5] and pluripotency [6]. Notwithstanding their pivotal role, these mechanisms are still poorly understood.

The epigenetic patterning of chromatin – the fibre made of DNA wrapped around histone proteins [1] – has been shown to strongly correlate with the three-dimensional (3D) nuclear organisation of interphase chromosomes [7–13]. For instance, transcriptionally active regions are often co-localised in super-enhancer hubs [8] or transcription factories [14], whereas transcriptionally inactive regions may form large heterochromatic *foci* [15], mega-base (Mbp) size lamin-associated domains [16] or Barr [17] and Polycomb bodies [18].

On the other hand, it is important to realise that the establishment of epigenetic patterns is fundamentally a *dynamic* process, where biochemical tags are constantly deposited, removed and degraded on histones, which can themselves be displaced during transcription or replaced after replication [19–23]. For this reason, simple models where epigenetic marks are stably deposited along chromatin [7, 9–11, 24, 25] are only crude approximations of a much more complex and dynamic scenario. Crucially, these "static" models fail to address key questions such as, how epigenetic patterns are first established along chromosomes, and how these change, for instance, with cellular ageing [5] or during disruptive events in the cell cycle [20]. In addition, an understanding of cell-to-cell variability in genome organisation [26, 27] and of efficiency of the induced-pluripotency pathway [28] may be achieved only through models which can take into account the plasticity of the epigenetic landscape.

In recent years, wide-spread technological advances in the field of molecular biology allowed the biophysical community to identify some of the key players in the dynamics of epigenetic patterning [5, 20, 29, 30]. At the heart of this process are proteins which can "read" and "write" biochemical tags along chromatin [20, 23, 31]. Importantly, some of these proteins can be found in the same complex [20] or are known to directly recruit one another [32]. For instance, the heterochromatin binding protein HP1 (a reader) possesses a chromodomain recognising tri-methylation of Lysine 9 on histone 3 (H3K9me3) [33–35]; at the same time, it can recruit the methyltransferase complex SUV39h1 [20] (a writer), therefore engaging a positive feedback loop which self-perpetuates this repressive mark [32, 36, 37].

Computational models of this positive feedback loop in 1D [38, 39] have shown that accounting for long-range contacts is necessary to allow spreading of repressive marks. On the other hand, 3D models coupling epigenetic and chromatin dynamics [40] have shown that spreading of silencing marks on a mobile 3D fibre can be viewed as an "all-or-none" transition, where a local fluctuation triggers an epigenetic wave of repressive marks which stabilises a compact globular state [40].

Existing experimental and computational studies suggest that the feedback between 3D chromatin structure and epigenetic dynamics along the chromosome may be a key potential mechanism underlying the establishment of epigenetic patterns which can display memory of their landscape. However, direct experimental observations of dynamical epigenetic and chromatin conformational changes in the nucleus are largely missing, in view of the difficulty to achieve enough spatio-temporal resolution *in vivo*. In light of this, here we propose, and investigate *in silico*, a novel avenue to directly observe the coupling between epigenetic dynamics and 3D chromatin folding.

The system we consider can be recreated *in vitro* via single-molecule experiments: it is inspired both by experiments studying the stretching of reconstituted chromatin [41–43], and by investigations of protein-DNA and protein-chromatin interactions *in vitro* [44, 45]. Specifically, we envisage a set-up where reconstituted chromatin is attached at both ends to macroscopic beads so as to exert stretching forces on the strand via optical or magnetic tweezers [41, 43]. Further, we imagine that the chromatin fibre is embedded in a solution where reader and writer proteins are added and activated. In practice, one may achieve this by including "readers" such as heterochromatin HP1 proteins [46] (which are known to be multivalent, hence act as bridges which fold chromatin), and "writers" such as SUV39h [20]. These proteins can respectively read and write the repressive H3K9me3 mark [32, 46]. Another option is to use Polycomb PRC complexes and Enhancer-of-zeste (E(z)) proteins which are respectively able to bind and deposit the H3K27me3 mark on histones [20, 30, 47–49].

Our results show that by tuning the strength of the external stretching force f and the binding affinity of reader proteins ϵ, the system can display either a stretched, epigenetically disordered, state (SD) or a compact, epigenetically ordered, one (CO). The two regimes are separated by an abrupt transition line which we quantitatively locate in the (f, ϵ) parameter space. We further show that when knots are tied along the chromatin, these can readily be identified by looking at the local order of epigenetic patterning. We argue that the topology of the knots localise and enhance 3D chromatin interactions, thereby "protecting" the epigenetically ordered region, which remains localised within the knotted arc. We dub these novel and remarkable states "epigenetic knotted solitons".

We finally envisage that our findings will inform the design of novel single-molecule experiments *in vitro* and illuminate the interplay between epigenetic dynamics and chromatin topology *in vivo*.

Chromatin Model and "Experimental" Set-up

We model the chromatin as a "recolourable" bead-spring polymer (for further details see Methods). Each bead has a size σ – for definiteness, we take $\sigma = 3$ kbp or 30 nm, but our results do not depend on this precise choice – and it bears a "colour" representing a specific biochemical tag. For instance, one may think of blue beads as representing chromatin regions with an excess of heterochromatin mark, H3K9me3, and red beads as ones with excess of Polycomb mark, H3K27me3. We further include grey beads which represent unmarked regions of chromatin. Beads interact through a Weeks-Chandler-Andersen potential, which strongly suppresses any overlap between beads. This interaction is further modified to account for attractive interactions (with affinity ϵ) between beads bearing the same colour. This effectively mimics the action of reading, or bridging, proteins which can bring together distant chromatin segments bearing the same epigenetic mark. Unmarked (grey) or differently marked beads are therefore considered to be solely sterically interacting. The potential can be written as depending on the colour (q) of interacting beads (a and b) as follows

$$U_{LJ}^{ab}(x) = \frac{4\epsilon_{ab}}{\mathcal{N}}\left[\left(\frac{\sigma}{x}\right)^{12} - \left(\frac{\sigma}{x}\right)^{6} - \left(\frac{\sigma}{x_c^{q_a q_b}}\right)^{12} + \left(\frac{\sigma}{x_c^{q_a q_b}}\right)^{6}\right] \tag{1}$$

for $x \leq x_c^{q_a q_b}$, whereas $U_{LJ}^{ab}(x) = 0$ otherwise. The potential can display an attractive (negative) region by setting $x_c^{q_a q_b} = 1.8\sigma$ for $q_a = q_b \neq 0$. Otherwise, when $q_a \neq q_b$ or one of the two beads is unmarked ($q = 0$) the cut-off is set to the minimum of the potential, i.e. $2^{1/6}\sigma$, yielding pure repulsion. The shift (last two terms in eq. (1)) ensure that there is no discontinuity in the potential. The binding affinity $\epsilon_{ab} = \epsilon$ when $q_a = q_b \neq 0$ and is a free parameter of our model, otherwise $\epsilon_{ab} = k_B T_L$ when $q_a \neq q_b$ or one of the two is unmarked.

The "recolouring" process is modelled via a Monte Carlo procedure which occurs at an inverse rate of $k_R^{-1} = \tau_R = 10^3 \tau_{Br} \simeq 10$ s (see Methods). Every time a bead (a) is selected for a recolouring attempt, its colour is randomly changed into one of the two remaining colours and the new energy is computed. Because the only colour-dependent potential employed in the simulations is the one written in Eq. (1), the difference in energy is simply given by

$$\Delta U = \sum_b U_{LJ}^{ab}(x)\,. \tag{2}$$

The recolouring attempt is then accepted with a Metropolis probability

$$p = \exp\left(\frac{-\Delta U}{k_B T_R}\right), \tag{3}$$

where T_R is a generic temperature that can, in principle, be tuned according to the efficiency of the writing process and is therefore independent on the solution temperature T_L, which regulates the stochastic dynamics of the beads in 3D [40] (see Materials and Methods). In practice, the chromatin strand can be viewed as a polymer whose segments display a further degree of freedom and it is therefore

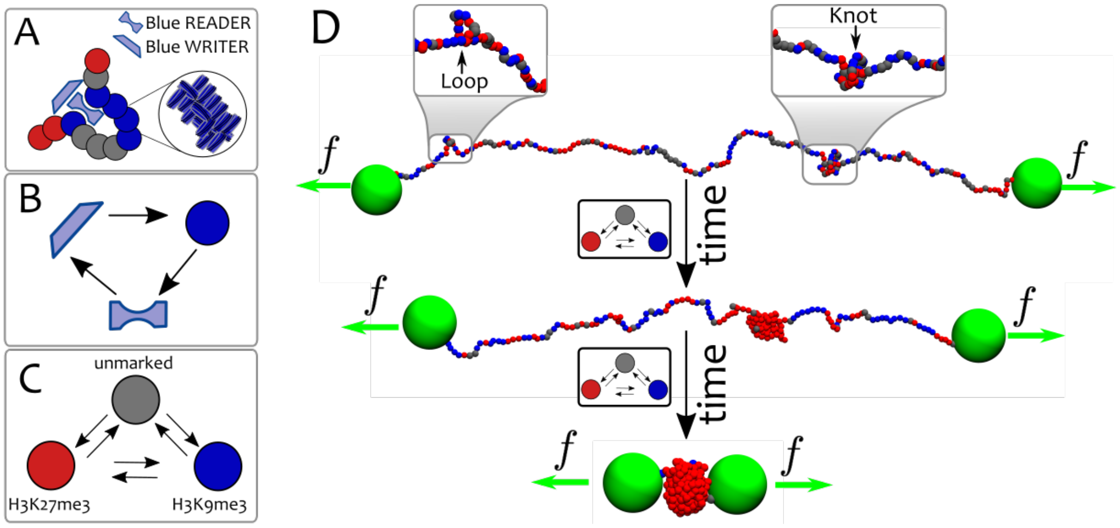

Figure 1. **Model and set-up.** (**A**) Chromatin is modelled as a coarse-grained "recolourable" bead-spring polymer; each bead represents few nucleosomes and it is coloured either red, blue or grey (unmarked bead). We implicitly model reader and writer proteins interacting with chromatin: the former allow polymer folding by bridging between segments bearing the same colour, the latter deposit new marks on nucleosomes and allow beads to change colour. (**B**) As red/blue marks are deposited on beads, respective reader proteins can bind to them; in turn, they recruit writer proteins, which deposit new red/blue marks. This mutual recruitment triggers a positive feedback loop. (**C**) Reader and writer proteins for red and blue marks compete over the same chromatin segments thereby generating dynamic epigenetic marking. (**D**) Experimental set-up: macroscopic beads (green) are attached at both chromatin ends and allow the strand to be stretched by external tweezers with force f [41, 43]. Starting from a configuration in equilibrium with the external force and with no proteins in solution, the system is flooded with reader/writer proteins thereby triggering the competition between red and blue marks (in the model, this is effectively done by allowing attractive interactions between chromatin segments). The system may eventually evolve towards an equilibrium state where chromatin is compact and only one epigenetic mark dominates ($\epsilon > \epsilon_c(f)$, shown in the figure), or towards one where chromatin is stretched and no coherent mark is established ($\epsilon < \epsilon_c(f)$, not shown).

related to a 1D chain of Ising (or Potts) spins allowed to move in 3D, also referred to as a "magnetic polymer" [50].

Here, we study the behaviour of the system upon tuning the stretching force f and the attraction strength ϵ, since this protocol may be realised *in vitro* by using reconstituted chromatin [41, 43] and proteins such as HP1 and SUV39 [36, 37], as previously mentioned. For simplicity, we limit to the case $T_L = T_R$ which ensures that the epigenetic read-write mechanism and the chromatin folding are governed by transition rules between different microstates that obey detailed balance and that can be described in terms of an effective free energy. Considering out-of-equilibrium conditions [40] leads to transitions between states with similar epigenetic patterns, i.e. from swollen ordered (or disordered) to compact ordered (or disordered), which can be understood as the homopolymer (or heteropolymer) limit of our system. Because these transitions do not shed light onto the interplay between epigenetics and chromatin conformations, we here decide not to pursue them.

In order to quantify the behaviour of the system we perform Brownian Dynamics simulations for a typical runtime of 5 10^6 τ_{Br} – corresponding to 1000 τ_R, or Monte Carlo sweeps – of a chain $L = 1000 \simeq 6$ Mbp beads long (if not specified otherwise). Although chromatin has not yet been reconstituted to such a length, the trends we uncover are generic and also hold for smaller values of L.

RESULTS

Epigenetic and Conformational Transitions of Stretched Chromatin

We initialise the systems by assuming that no reader or writer protein is present at $t < 0$ and by letting the chromatin equilibrate while subject to the stretching force f and in a bath at temperature T_L. In practice, we do this by lowering the cut-off for the Weeks-Chandler-Andersen potential written in Eq. (1) to $2^{1/6}\sigma$ for any pair of beads so that the interactions between beads are purely repulsive.

At time $t = 0$ the system is flooded with reader and writer proteins which can bridge beads with the same epigenetic mark (i.e., colour) and to attempt recolouring of the beads (i.e., writing) at a rate τ_R^{-1}. To visualise the epigenetic evolution of the system, in Fig. **2**(A-B) we report typical "kymographs", which show the colour (q) of each bead along the polymer at a given time t. Kymographs readily capture the epigenetic "ordering" of the chromatin strand as a function of time for a specific choice of ϵ and

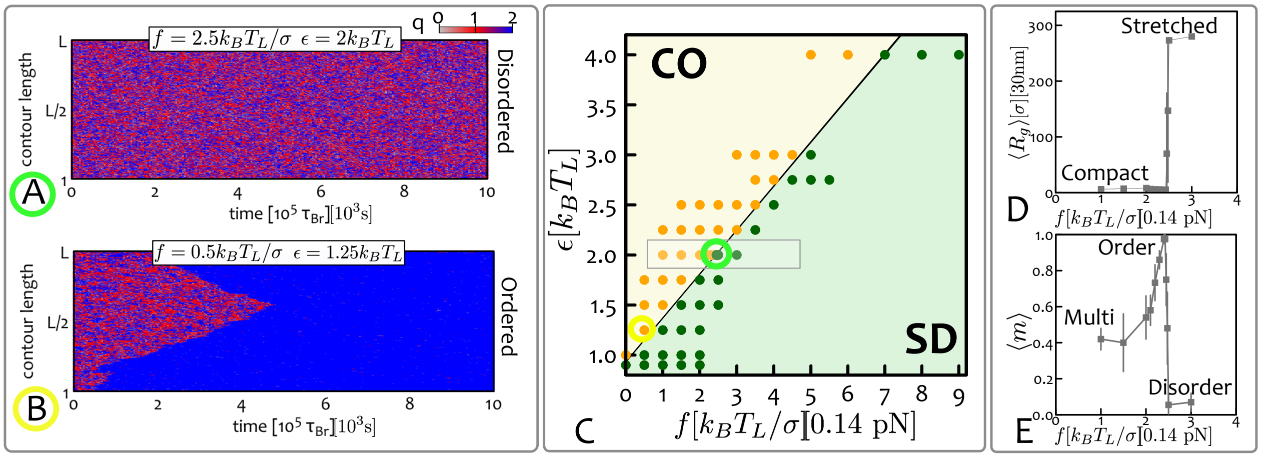

Figure 2. **Epigenetically-Driven Phase Transition in Stretched Chromatin**. At time $t = 0$ the chromatin fibre $L = 1000$ beads long is in equilibrium, under a given stretching force f. The system is then instantaneously flooded with (implicitly modelled) reader and writer proteins, and the attractive interaction between same-coloured beads turned on. (**A**-**B**) The evolution of the system as a function of time can be visualised through "kymographs" which show the colour (q) of each chromatin segment (y-axis) as a function of time (x-axis). Depending on the choice of parameters ϵ and f, the system may display (**A**) stretched-"epigenetically disordered" (SD) states or (**B**) compact-"epigenetically ordered" (CO) ones. From the kymographs, one can also appreciate the nucleation and spreading dynamics of epigenetic marks. (**C**) The phase diagram of the system in the parameter space (f, ϵ) displays two regions with compact-ordered (CO, yellow shaded) and stretched-disordered (SD, green shaded) equilibrium states. The regions are separated by an abrupt transition line $f_c(\epsilon)$. Data-points corresponding to performed simulations are also shown. (**D**-**E**) Show the force dependent of the mean radius of gyration $\langle R_g \rangle$ and the mean absolute magnetisation $\langle m \rangle$ for a fixed value of ϵ. Both profiles display an abrupt transition when the critical line $f_c(\epsilon)$ is crossed. Intriguingly, the absolute magnetisation decreases as $f \to 0$, which can be understood in terms of multiple nucleation points that trigger meta-stable multi-domain states (see below). Data-points for $\epsilon = 2k_BT_L$ and near the transition line have been obtained by averaging over 64 independent replicas. The grey rectangle in **(C)** highlights the region considered for the profiles reported in panels **(D-E)**. Mapping to real units of measure are also shown. See also supplementary Movies.

f. These two latter quantities are the main free parameters of the system: the stretching force f can be controlled *in vitro* through either optical or magnetic tweezers, while the binding affinity between readers and epigenetic marks may be varied by considering mutants of bridging proteins such as HP1 [36, 46].

In order to further quantify the behaviour of the system in parameter space, we systematically vary f and ϵ and quantify the equilibrium states by measuring the values of the absolute "epigenetic magnetisation"

$$m \equiv \frac{|n_{\text{red}} - n_{\text{blue}}|}{L} \tag{4}$$

and of the average chromatin extension via its radius of gyration

$$R_g^2 = \frac{1}{L}\sum_{i=0}^{L} [\boldsymbol{r}_i - \boldsymbol{r}_{\text{CM}}]^2 \ , \tag{5}$$

where $\boldsymbol{r}_i$ and $\boldsymbol{r}_{\text{CM}}$ are the positions of segment i and of the centre of mass of the chain, respectively. In order to construct an equilibrium phase diagram of the system, we estimate the average $\langle R_g \rangle \equiv \langle R_g^2 \rangle^{1/2}$ and $\langle m \rangle$, first by time-averaging over a trajectory at steady state and then averaging the results across several trajectories. By measuring these two observables for different choices of ϵ and f we report the phase diagram shown Fig. **2**(C). One can readily notice that the typical equilibrium configurations can be grouped into two distinct phases: one is compact and epigenetically ordered (CO), i.e. with

$$\langle R_g \rangle \sim L^{1/3}\sigma, \ \langle m \rangle > 0 \tag{6}$$

whereas the other is stretched and epigenetically disordered (SD), i.e. with

$$\langle R_g \rangle \sim L\sigma, \ \langle m \rangle \simeq 0 \, . \tag{7}$$

These two regimes are separated by a transition line and one can readily appreciate from the profiles of $\langle R_g \rangle$ and $\langle m \rangle$ (Figs. **2**(D-E)) that this transition is abrupt. In the limit of stretching forces $f \to 0$, we retrieve the first-order transition observed in Ref. [40] for un-stretched chromatin. When larger stretching forces are applied, the system therefore displays a force-dependent critical line $\epsilon_c(f)$ (or $f_c(\epsilon)$) which retains the first-order-like features observed at $f = 0$.

Within the CO region, there is a parameter range where a multi-domain epigenetic structure emerges dynamically. This can be readily seen from the profile of the epigenetic magnetisation $\langle m \rangle$ in Fig. **2**(E): near the transition line (from the compact ordered side) this quantity displays a sharp peak which then decreases for smaller values of the

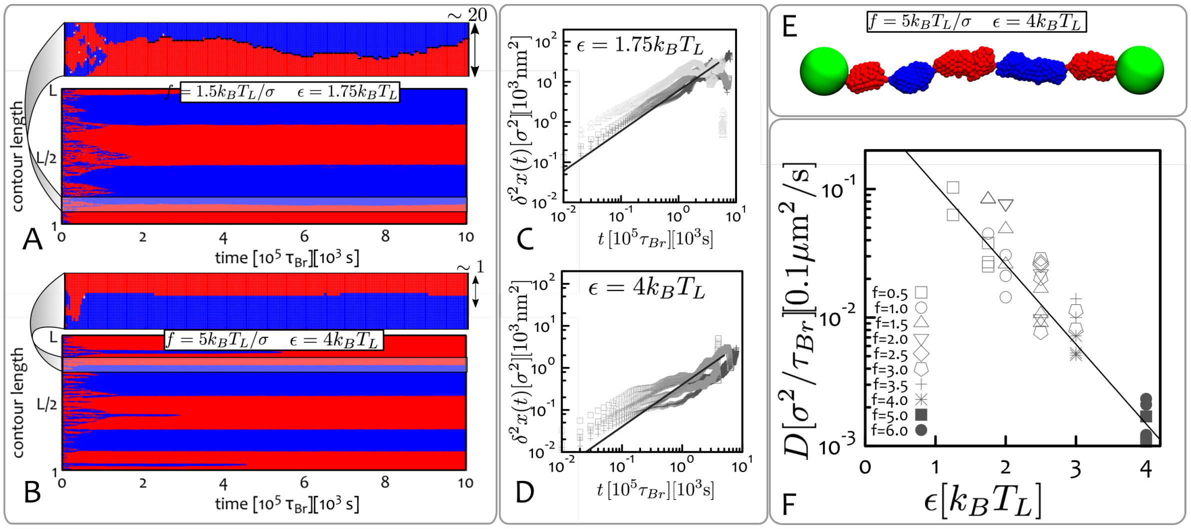

Figure 3. **Epigenetic Domains and Boundary Diffusion.** (**A-B**) These panels show kymographs of systems displaying long-lived multi-domain states. Although these are not stable (i.e., free-energy minimising) states, they evolve on time-scales much longer than the simulations (or experiment) runtime. Zoomed in regions displayed at the top of each kymograph pinpoint the dynamics of boundaries between epigenetic domains. (**C-D**) These panels show typical mean-squared displacements of the boundaries $\delta^2x(t)$ obtained from the kymographs over multiple system replicas and for a given ϵ. Symbols are grouped corresponding to the applied stretching force f. (**E**) Typical snapshot of the system whose kymograph is reported in (**B**), i.e. $\epsilon = 4k_BT_L$ and $f = 5k_BT_L/\sigma$. (**F**) Plot of the diffusion coefficient D extracted from $\delta^2x(t)$. Data-points are grouped into symbols corresponding to equal stretching force f (in units of k_BT_L/σ) and plotted as a function of ϵ in a log-linear plot to highlight the exponential decay. Units of measure mapped to real length- and time-scales are shown. See also Supplementary movies.

force. In the region labelled as "multi" in Fig. **2**(E), the magnetisation is lower than unity because multiple macroscopic ordered regions populated by different epigenetic marks compete with one another, thereby lowering the overall magnetisation. These domains must be metastable (though long-lived) as in steady state a single domain is preferred since it removes domain walls, which have a free energy cost. The existence of multi-domain patterns arises because, at stretching forces far from critical $f_c(\epsilon)$, there may be multiple nucleation points along the fibre for the spreading of epigenetic marks. These nucleation points are generated by a local increase in chromatin density, for instance through the transient formation of loops and 3D interactions, which then trigger local spreading of different epigenetic marks (see Fig. **3**). As the stretching force gets closer to the critical value $f_c(\epsilon)$, the nucleation probability declines, hence only one epigenetic mark is able to take over the whole chromatin strand, thereby enhancing its overall epigenetic magnetisation.

It is useful to compare the epigenetic transition between the CO and the SD state in Fig. **2** with the equilibrium transition between a compact and a stretched state which can be observed in a stretched homopolymer with self-attractive interactions (i.e., in a poor solvent). While the transition is first order in the homopolymer case as well [51, 52], here the epigenetic degrees of freedom increase the entropy of the disordered phase, thereby the value of the critical force, for a given ϵ, is lower. Another key difference is that multi-pearl structures are only observed transiently close to the transition for the homopolymer case [51], whereas in the epigenetic case multi-domain states arise far from the transition, and they are long-lived. Both these differences should be experimentally detectable, as the homopolymer case may be recreated by reconstituting chromatin fibres with histone octamers with controlled biochemical tags or in the presence of linker histones (H1), while avoiding the action of writers.

Long-Lived Epigenetic Domains and Boundary Diffusion

The competition of epigenetic marks over a chromatin segment is a phenomenon relevant to many biological systems. For example, genes that are positioned near telomeres may switch between transcriptionally active and silenced states from one generation to the next [55]. This phenomenon, generally referred to as "telomeric position effect", is particularly relevant in yeast [55] but it has also been observed in human cells [56]. It occurs because heterochromatin marks largely populating the telomeric gene-poor regions of the genome invade gene-rich stretches of chromatin, effectively switching off gene transcription.

In our proposed single molecule set-up, the emergence of multiple competing epigenetic domains in reconstituted chromatin under (small) tension allows a quantitative study of stability and diffusive dynamics of boundaries between neighbouring epigenetic domains, processes which

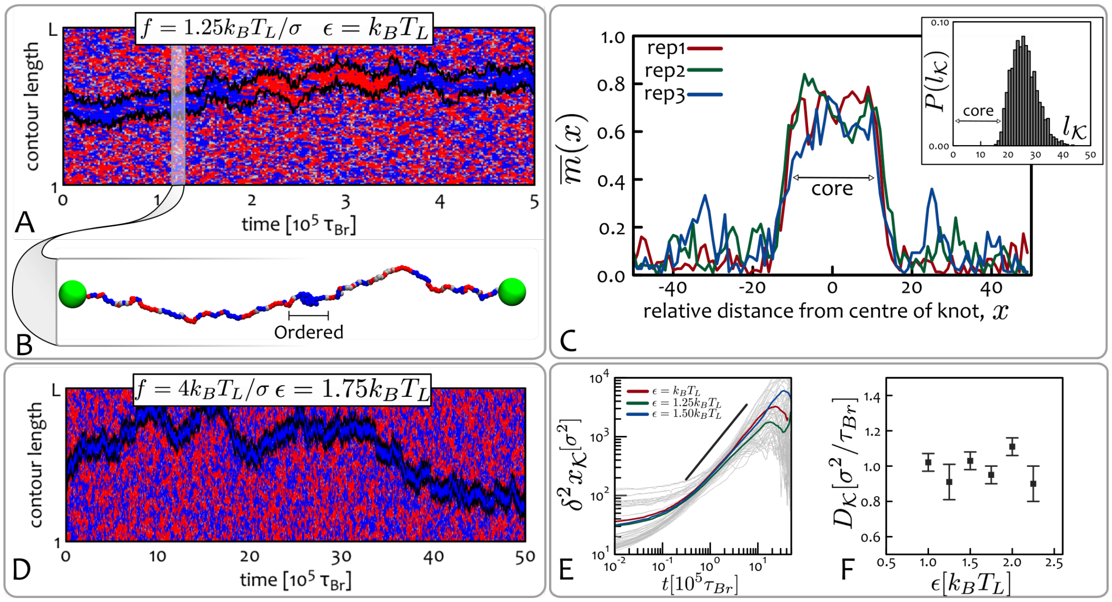

Figure 4. **Epigenetic knotted solitons.** (**A**) Kymograph corresponding to simulations initiated with a knotted chromatin fibre in equilibrium, with $f = 1.25k_BT_L/\sigma$. At time $t = 0$ the system is flooded with reader and writer proteins having binding affinity $\epsilon = kBT_L$. The chromatin contains a figure-of-eight (4_1) knot [53], whose escape from the fibre is avoided by the macroscopic beads at the terminal ends. Superimposed on the kymograph, we also show the boundaries of the knot computed from the 3D polymer configurations through the knot identification algorithm described in Ref. [54] (black lines). Intriguingly, from this plot one can immediately realise that the knotted portion of the chain can be identified with the epigenetically ordered region, i.e. the chromatin segment where the beads are homogeneously coloured. (**B**) Snapshot of the chromatin fibre at the time-step highlighted in the kymograph. (**C**) Analysis of the knotted soliton for different system replicas. In the main panel we show the time-averaged magnetisation $\overline{m}(x)$ at position x relative to the centre of the knot (see Materials and Methods); this displays a localised increase which we identify as an "epigenetic soliton". In the inset we show the probability of observing a knotted arc of length $l_\mathcal{K}$ through the knot identification algorithm of Ref. [54]. As one can notice, the length of the epigenetically ordered region and that of the knotted arc are in near-perfect agreement. This strongly suggests that the localised knotted segment leads to the local coherency in the epigenetic marks. **(D)** Increased affinity ϵ leads to slower switching times: here the epigenetic soliton never changes state during a 10 times longer simulation. **(E-F)** The measured mean squared displacement $\delta^2 x_\mathcal{K}(t)$ and diffusion coefficients $D_\mathcal{K}$ of the knotted solitons are shown to be insensitive to the interaction strength ϵ. Here we consider a chromatin fibre with $L = 200$ beads. See also Supplementary Movies.

may be relevant to understand the mechanisms underlying competition and spreading of epigenetic marks *in vivo*.

In practice, an epigenetic domain in the stretched chromatin fibre is defined as a macroscopic region (e.g., consisting of $w \geq 50$ beads) over which more than $th = 90\%$ of the beads are homogeneously coloured. A boundary between domains is found if the beads on either side of the boundary have opposite colours (i.e., red and blue) and the signed magnetisation difference Δm of the left and right domain is such that $|\Delta m| > 2th$ (i.e. the domains have opposite signed magnetisation).

By tracking the position of the domain boundaries over time, one can measure their mean square displacement

$$\delta^2 x(t) = \frac{1}{T-t}\int_0^{T-t} \left[x(t_0+t) - x(t_0)\right]^2 dt_0 \qquad (8)$$

where $x(t)$ is the position of a given boundary at time t. Plots of $\delta^2 x(t)$ for different replicas and two choices of interaction energies ϵ are reported in Fig. **3**(C-D). Data points are grouped according to the applied stretching force f, showing that the dependence on f is much weaker compared to that on ϵ. We extract the diffusion coefficient D as the long time behaviour of $\delta^2 x(t)$, i.e.

$$D = \lim_{t\to\infty} \frac{\delta^2 x(t)}{2t}\,, \qquad (9)$$

which is reported in Fig. **3**(F) and clearly shows an exponential decay of D as a function of ϵ.

In our stretched chromatin assay, the diffusive dynamics of epigenetic boundaries is therefore mainly controlled by the strength of the attractive interactions between beads bearing the same epigenetic mark. *In vivo*, another possible important factors affecting diffusivity of epigenetic boundaries may be the presence of insulators and architectural proteins such as CTCF and cohesins [1, 44, 45]. It would be interesting to adapt our model to account for these features in the future.

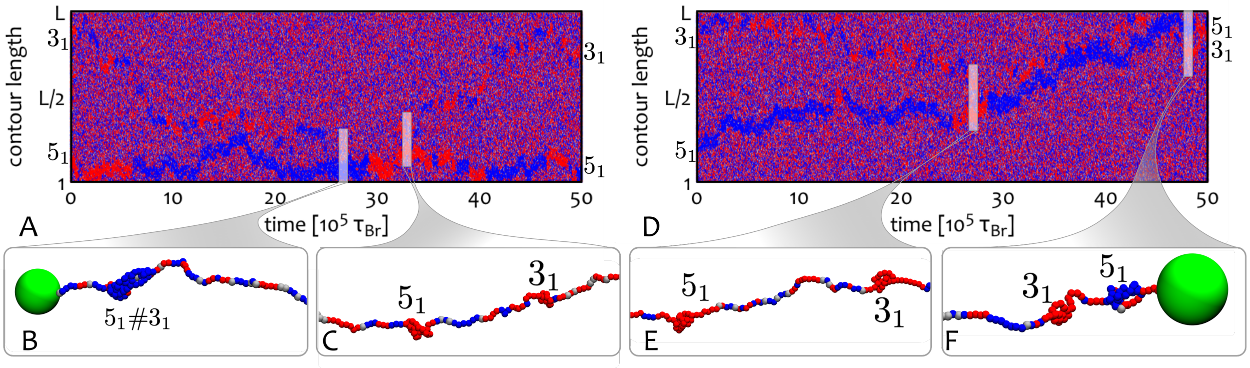

Figure 5. **Interacting Solitons.** The kymographs display reflecting **(A)** or cross-through [57] **(D)** interactions between epigenetic knotted solitons. Panels **(B-C)** and **(E-F)** show typical situations where knots can form a unique composite knot or form independent domains. Here the chromatin strand has length $L = 400$ beads.

Epigenetic Knotted Solitons

The abrupt, first order, phase transition between the compact-ordered and the stretched-disordered state shown in Figure **2** is here due to the coupling between *global* 3D structure and 1D epigenetic dynamics [40], which gives rise to a positive feedback loop where compaction aids spreading, which leads to further crumpling at small enough f, or sufficiently large ϵ.

It is interesting to ask whether one may design an experimentally realisable situation whereby this positive feedback is only realised *locally*, rather than globally as done in Ref. [40]. To do so, we consider a knotted chromatin fibre: the idea behind this construct is that a polymeric knot tightly localises upon stretching [58, 59], and that the knotted region is more compact, thereby providing a natural nucleation point for epigenetic ordering. In the single molecule experiment set-up we consider (Fig. **2**), a physical knot may be tied within chromatin by using, for instance, micro-manipulating techniques used to generate knots on DNA [60–62]. Alternatively, a knotted chromatin fibre may be self-assembled by first knotting the naked DNA strand, and later allowing nucleosome formation by adding histone octamers to the solution [63].

In Figure **4** (A-B) we show the kymograph for a chromatin strand along which a figure-of-eight (4_1) knot is tied. As one can see, even though we choose values of ϵ and f so that the system is in the stretched-disordered region of the phase diagram (see Fig. **2**), the kymograph clearly displays a localised epigenetically ordered region. To confirm whether this region corresponds to the knotted arc, tightened upon stretching and self-attraction, we pinpoint and monitor the time evolution of the knotted region by resorting to a top-down search of the smallest portion of chromatin that yields, upon suitable closure, a ring with the same topology of the whole chain [54]. Remarkably, the knotted portion found with this well-established algorithm perfectly matches the epigenetic domain in the kymograph (black lines in Fig. **4** identify the boundary of the knotted arc, see also Methods).

We call this remarkable configuration an "epigenetic knotted soliton", because, as a soliton, it travels along the fibre by keeping near-constant shape, hence displaying particle-like behaviour. Similarly to solitons recently observed in meta-materials [64] and minimal surfaces [65], it is the non-trivial topology of the localised knot which keeps the structure together. The epigenetic state of the knotted soliton can "switch" from red to blue, or vice versa – switching events occur, in Figure **4**, at about 2 $10^5\tau_{Br}$ and 3.5 $10^5\tau_{Br}$). Switching between the two stable states is observed because the soliton has a finite size, hence the effective free energy barrier separating the red and blue states can be occasionally overtaken by fluctuations.

The physical properties of the soliton are tunable: for instance, by increasing the stretching force the knot tightens and becomes more compact, so that the soliton shrinks in size but becomes more stable due to the increase of 3D interactions which locally "protect" the epigenetically ordered state. By increasing values of ϵ, such that the system remains overall disordered ($\epsilon < \epsilon_c(f)$) the switching rate decreases (i.e., the epigenetic ordering is more robust, see Fig. **4**(D)) – the diffusion coefficient, however, does not depend on ϵ appreciably (unlike for the case of epigenetic domains, see Fig. **3**(E-F)). It would be of interest to quantify how the soliton diffusion coefficient depends on knot type, and whether, as in knots in swollen polymers, twist knots and more complex knots are less mobile [53, 61, 66, 67].

For values of $\epsilon > \epsilon_c(f)$, we observe the same SD-CO phase transition reported in Figure **2**, where the localised knot now acts as a nucleation point for the spreading of its epigenetic mark.

Finally, we investigate how knotted solitons interact on a chromatin strand. We consider a system where two knots, a pentafoil 5_1 and a trefoil 3_1, are initially tied along the chromatin. As the system evolves each knot forms its independent epigenetic soliton. As shown by the kymographs

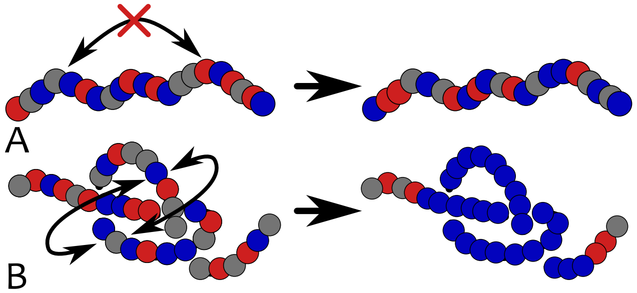

Figure 6. **Epigenetics and 3D topology.** This figure intuitively shows the reason behind the formation of a stable epigenetic domain within the knotted chromatin segment: the local 3D topology of a knot "protects" the underlying epigenetic information by enhancing local 3D interactions (**B**). On the contrary, an unknotted stretched segment lacks long-range interactions required to sustain epigenetic coherency (**A**).

in Figure **5**(A-D), both solitons diffuse along the chain [61]; they also interact and may merge to form a single domain, and later on split again (Fig. **5**(A)), or even cross each other [57] (Fig. **5**(D)). Following every collision or interaction, each soliton retains its overall structure and re-establishes its locally coherent epigenetic mark.

DISCUSSION AND CONCLUSIONS

In several areas of chromosome biology, researchers assume the existence of a tight coupling between 3D chromatin structure and epigenetic dynamics, as this coupling provides an appealing mechanism for the *de novo* establishement and maintainance of epigenetic patterns [38, 40, 46, 47, 49]. However, as both epigenetic spreading and chromatin dynamics may occur on comparable and relatively fast timescales *in vivo* (minutes to hours [20, 68]), it is extremely difficult to design an experiment to demonstrate this coupling dynamically inside the cell. Here, instead, we proposed, and investigated *in silico*, a set-up for an experiment to test and quantify the coupling directly *in vitro*. The set-up involves a reconstituted chromatin fibre [41, 43], which is stretched (e.g., by laser tweezers) and interacts with an ensemble of reader and writer proteins – such as heterochromatin HP1 and SUV39h1 complexes [36, 46] or polycomb [47, 49] enzymes able to bind to and deposit their respective marks.

We showed that by varying, for instance, the stretching force, one can trigger a phase transition between a compact-ordered phase, where a single epigenetic mark invades the whole chromatin fibre, and a stretched-disordered phase, where no single mark can take over the system. It is important that this transition may be observed by varying only the stretching force, as this is simpler to fine tune continuously *in vitro* with respect to temperature (which can inactivate readers or writers altogether) and effective self-attraction between chromatin segments (which may be achieved through the use of mutant reader proteins).

By using single-molecule imaging [44, 45, 69] or super-resolution techniques [70] one may envisage to employ antibodies recognising H3K9me3 and H3K27me3 epigenetic states to separately stain the marks and thereby discriminate between an epigenetically ordered and an epigenetically disordered state. In principle, one may even be able to observe the transition from one to the other by increasing the stretching force in the single molecule assay. Another possibility to assess the degree of epigenetic order is to use single-cell chromatin immunoprecipitation [71] (ChIP) on the reconstituted chromatin, to quantify histone modifications along the fibres. Note that, for this approach to be viable, one would have to use genomic DNA to reconstitute chromatin (rather than repeating, such as 601, sequences), so that locations along the DNA can be mapped uniquely. Therefore, whilst an experimental investigation of the compact-ordered to stretched-disordered transition is certainly challenging, and will require state-of-the-art experimental techniques, it is in principle feasible, and would constitute the first direct measurement of the so far elusive coupling between chromatin structure and epigenetic dynamics.

By tying a knot along the reconstituted chromatin fibre we observed that the system can harbour a novel structure, an "epigenetically knotted soliton" (see Figs. **4**- **6**). This is a tight knot which is locally ordered epigenetically and diffuses freely, within an epigenetically disordered background. The size of the soliton is tunable by varying the stretching force applied to the chromatin fibre, and different solitons interact in a variety of ways: they may bounce off one another or cross through each other when they collide. We also expect similar topological solitons to be universally found in knotted magnetic polymer, a new kind of topological soft matter which has not yet been realised experimentally [50].

The findings we have reported here may also be employed to *detect* knots in chromatin strands. Starting from an ordered and crumpled state (such as heterochromatin [1]), one may imagine to apply and continuously increase an external stretching force f and monitor the evolution of the system; whereas ordered unknotted regions will undergo an abrupt phase transition and become epigenetically disordered (see phase diagram in Fig. **2**), knotted chromatin will instead preserve a localised epigenetically ordered region that can be identified with the knotted segment (Fig. **4**).

Besides all this, the epigenetic solitons provide another observable consequence of the dynamic coupling between 3D chromatin structure and epigenetics which can be tested in future single-molecule experiments.

Acknowledgements

We acknowledge ERC for funding (Consolidator Grant THREEDCELLPHYSICS, Ref. 648050).

---

[1] B. Alberts, A. Johnson, J. Lewis, D. Morgan, and M. Raff, *Molecular Biology of the Cell* (Taylor & Francis, 2014) p. 1464.
[2] C. H. Waddington, Nature **218**, 639 (1968).
[3] C. H. Waddington, Nature **150**, 563 (1942).
[4] N. Naumova, M. Imakaev, G. Fudenberg, Y. Zhan, B. R. Lajoie, L. A. Mirny, and J. Dekker, Science (80-.). **342**, 948 (2013).
[5] S. Pal and J. Tyler, Sci. Adv. **2**, 253 (2016).
[6] S. Yamanaka and H. M. Blau, Nature **465**, 704 (2010).
[7] S. S. P. Rao, M. H. Huntley, N. C. Durand, E. K. Stamenova, I. D. Bochkov, J. T. Robinson, A. L. Sanborn, I. Machol, A. D. Omer, E. S. Lander, and E. L. Aiden, Cell **159**, 1665 (2014).
[8] R. A. Beagrie, A. Scialdone, M. Schueler, D. C. A. Kraemer, M. Chotalia, S. Q. Xie, M. Barbieri, I. de Santiago, L.-M. Lavitas, M. R. Branco, J. Fraser, J. Dostie, L. Game, N. Dillon, P. A. W. Edwards, M. Nicodemi, and A. Pombo, Nature **543**, 519 (2017).
[9] M. Barbieri, M. Chotalia, J. Fraser, L.-M. Lavitas, J. Dostie, A. Pombo, and M. Nicodemi, Proc. Natl. Acad. Sci. USA **109**, 16173 (2012).
[10] C. A. Brackley, D. Michieletto, F. Mouvet, J. Johnson, S. Kelly, P. R. Cook, and D. Marenduzzo, Nucleus **7**, 453 (2016).
[11] D. Jost, Phys. Rev. E **89**, 1 (2014).
[12] J. R. Dixon, S. Selvaraj, F. Yue, A. Kim, Y. Li, Y. Shen, M. Hu, J. S. Liu, and B. Ren, Nature **485**, 376 (2012).
[13] C. A. Brackley, J. Johnson, S. Kelly, P. R. Cook, and D. Marenduzzo, Nucleic Acids Res. (2016).
[14] P. Cook, *Principles of Nuclear Structure and Function* (Wiley, 2001).
[15] N. Gilbert, S. Gilchrist, and W. a. Bickmore, Int. Rev. Cytol. **242**, 283 (2004).
[16] L. Guelen, L. Pagie, E. Brasset, W. Meuleman, M. B. Faza, W. Talhout, B. H. Eussen, A. de Klein, L. Wessels, W. de Laat, and B. van Steensel, Nature **453**, 948 (2008).
[17] S. F. Pinter, R. I. Sadreyev, E. Yildirim, Y. Jeon, T. K. Ohsumi, M. Borowsky, and J. T. Lee, Genome Res. **22**, 1864 (2012).
[18] A. H. Wani, A. N. Boettiger, P. Schorderet, A. Ergun, C. Münger, R. I. Sadreyev, X. Zhuang, R. E. Kingston, and N. J. Francis, Nat. Commun. **7**, 10291 (2016).
[19] S. Ramachandran and S. Henikoff, Sci. Adv. **1**, e1500587 (2015).
[20] G. E. Zentner and S. Henikoff, Nat. Struct. Mol. Biol. **20**, 259 (2013).
[21] A. V. Probst, E. Dunleavy, and G. Almouzni, Nat. Rev. Mol. Cell. Biol. **10**, 192 (2009).
[22] N. Saksouk, E. Simboeck, and J. Déjardin, Epigenetics Chromatin **8**, 3 (2015).
[23] T. Zhang, S. Cooper, and N. Brockdorff, EMBO Rep **16**, 1467 (2015).
[24] T. Sexton, E. Yaffe, E. Kenigsberg, F. Bantignies, B. Leblanc, M. Hoichman, H. Parrinello, A. Tanay, and G. Cavalli, Cell **148**, 458 (2012).
[25] G. J. Filion, J. G. van Bemmel, U. Braunschweig, W. Talhout, J. Kind, L. D. Ward, W. Brugman, I. J. de Castro, R. M. Kerkhoven, H. J. Bussemaker, and B. van Steensel, Cell **143**, 212 (2010).
[26] T. Nagano, Y. Lubling, T. J. Stevens, S. Schoenfelder, E. Yaffe, W. Dean, E. D. Laue, A. Tanay, and P. Fraser, Nature **502**, 59 (2013).
[27] T. Nagano, Y. Lubling, C. Varnai, C. Dudley, W. Leung, Y. Baran, N. Mandelson Cohen, S. Wingett, P. Fraser, and A. Tanay, bioRxiv (2016), 10.1101/094466.
[28] S. Yamanaka, Cell Stem Cell **1**, 39 (2007).
[29] S. Henikoff and J. M. Greally, Curr. Biol. **26**, R644 (2016).
[30] P. B. Talbert and S. Henikoff, Nat. Rev. Genet. **7**, 793 (2006).
[31] C. Willyard, Nature **542**, 406 (2017).
[32] M. M. Müller, B. Fierz, L. Bittova, G. Liszczak, and T. W. Muir, Nat. Chem. Biol. **12**, 188 (2016).
[33] M. Lachner, D. O'Carroll, S. Rea, K. Mechtler, and T. Jenuwein, Nature **410**, 116 (2001).
[34] K. Hiragami-Hamada, S. Soeroes, M. Nikolov, B. Wilkins, S. Kreuz, C. Chen, I. A. De La Rosa-Velázquez, H. M. Zenn, N. Kost, W. Pohl, A. Chernev, D. Schwarzer, T. Jenuwein, M. Lorincz, B. Zimmermann, P. J. Walla, H. Neumann, T. Baubec, H. Urlaub, and W. Fischle, Nat. Commun. **7**, 11310 (2016).
[35] S. Kilic, A. L. Bachmann, L. C. Bryan, and B. Fierz, Nat. Commun. **6**, 7313 (2015).
[36] D. Canzio, A. Larson, and G. J. Narlikar, Trends Cell Biol. **24**, 377 (2014).
[37] B. Al-Sady, H. D. Madhani, and G. J. Narlikar, Mol. Cell **51**, 80 (2013).
[38] I. B. Dodd, M. a. Micheelsen, K. Sneppen, and G. Thon, Cell **129**, 813 (2007).
[39] I. B. Dodd and K. Sneppen, J. Mol. Biol. **414**, 624 (2011).
[40] D. Michieletto, E. Orlandini, and D. Marenduzzo, Phys. Rev. X , 1 (2016).
[41] M. Kruithof, F.-T. Chien, A. Routh, C. Logie, D. Rhodes, and J. van Noort, Nat. Struct. Mol. Biol. **16**, 534 (2009).
[42] H. Meng, K. Andresen, and J. Van Noort, Nucleic Acids Res. **43**, 3578 (2015).
[43] Y. Cui and C. Bustamante, Proc. Nat. Acad. Sci. USA. **97**, 127 (2000).
[44] M. Kanke, E. Tahara, P. J. Huis, and T. Nishiyama, EMBO J. , 1 (2016).
[45] J. Stigler, G. Ö. Çamdere, D. E. Koshland, and E. C. Greene, Cell Rep. **15**, 988 (2016).
[46] D. Canzio, M. Liao, N. Naber, E. Pate, A. Larson, S. Wu, D. B. Marina, J. F. Garcia, H. D. Madhani, R. Cooke, P. Schuck, Y. Cheng, and G. J. Narlikar, Nature **496**, 377 (2013).
[47] A. Angel, J. Song, C. Dean, and M. Howard, Nature **476**, 105 (2011).
[48] F. Laprell, K. Finkl, and J. Müller, Science (80-.). **8266**, eaai8266 (2017).
[49] F. Ciabrelli, F. Comoglio, S. Fellous, B. Bonev, M. Ninova, Q. Szabo, A. Xuéreb, C. Klopp, A. Aravin, R. Paro, F. Bantignies, and G. Cavalli, Nat. Genet. (2017), 10.1038/ng.3848.
[50] T. Garel, H. Orland, and E. Orlandini, EPJ B **268**, 261 (1999).
[51] P. Grassberger and H.-P. Hsu, Phys. Rev. E **65**, 031807 (2002).

[52] D. Marenduzzo, A. Maritan, A. Rosa, and F. Seno, Phys. Rev. Lett. **90**, 088301 (2003).
[53] C. C. Adams, *The knot book: An elementary introduction to the mathematical theory of knots* (W H Freeman and Company (New York), 1994).
[54] L. Tubiana, E. Orlandini, and C. Micheletti, Prog. Theor. Phys. Suppl. **191**, 192 (2011).
[55] D. E. Gottschling, O. M. Aparicio, B. L. Billington, and V. A. Zakian, Cell **63**, 751 (1990).
[56] J. A. Baur, Y. Zou, J. W. Shay, and W. E. Wright, Science (80-.). **292**, 2075 (2001).
[57] B. Trefz, J. Siebert, and P. Virnau, **111**, 7948 (2014).
[58] O. Farago, Y. Kantor, and M. Kardar, EPL (Europhysics Letters) **60**, 53 (2002).
[59] M. Caraglio, C. Micheletti, and E. Orlandini, Physical review letters **115**, 188301 (2015).
[60] Y. Arai, R. Yasuda, K.-i. Akashi, Y. Harada, H. Miyata, K. Kinosita, and H. Itoh, Nature **399**, 446 (1999).
[61] X. R. Bao, H. J. Lee, and S. R. Quake, Phys. Rev. Lett. **91**, 265506 (2003).
[62] D. Liu, G. Chen, U. Akhter, T. M. Cronin, and Y. Weizmann, Nature Chemistry (2016).
[63] C. A. Brackley, J. Allan, D. Keszenman-Pereyra, and D. Marenduzzo, Nucleic Acids Res. **43**, 63 (2015).
[64] B. G.-g. Chen, N. Upadhyaya, and V. Vitelli, Proc. Nat. Acad. Sci. USA **111**, 13004 (2014).
[65] T. Machon, G. P. Alexander, R. E. Goldstein, and A. I. Pesci, Phys. Rev. Lett. **017801**, 1 (2016).
[66] M. Di Stefano, L. Tubiana, M. Di Ventra, and C. Micheletti, Soft Matter **10**, 6491 (2014).
[67] R. Matthews, A. Louis, and J. Yeomans, EPL (Europhysics Letters) **89**, 20001 (2010).
[68] T. K. Barth and A. Imhof, Trends Biochem. Sci. **35**, 618 (2010).
[69] H. Yardimci, A. B. Loveland, A. M. van Oijen, and J. C. Walter, Methods **57**, 179 (2012).
[70] B. Huang, W. Wang, M. Bates, and X. Zhuang, Science **319**, 810 (2008).
[71] A. Rotem, O. Ram, N. Shoresh, R. A. Sperling, A. Goren, D. A. Weitz, and B. E. Bernstein, Nat. Biotechnol. **33**, 1165 (2015).
[72] K. Kremer and G. S. Grest, J. Chem. Phys. **92**, 5057 (1990).

Material and Methods

Chromatin Model

Chromatin is modelled as a bead-spring polymer chain [72] where each bead has nominal size σ. Harmonic springs between beads a and b are imposed as $U_H^{ab}(x) = (\delta_{b,a+1} + \delta_{b,a-1})k_BT_L\kappa(x-x_0)/2$ with $x_0 = 1.1\sigma$ and $\kappa = 100k_BT_L$ to ensure the connectivity of the backbone. The chain stiffness is regulated by a Kratky-Porod potential between triplets of beads forming an angle $\theta = (\boldsymbol{t}_a \cdot \boldsymbol{t}_b)/(|\boldsymbol{t}_a||\boldsymbol{t}_b|)$ where $\boldsymbol{t}_a$ is the vector joining beads a and $a+1$ as $U_{KP}^{ab}(\theta) = (\delta_{b,a+1} + \delta_{b,a-1})k_BT_Ll_p(1+\cos\theta)/\sigma$ with $l_p = 3\sigma$. The total potential $U^a(x)$ experienced by each bead is given by the sum over all the possible interacting pairs and triplets, i.e.

$$Ua(x) = \sum_b \left[U_{LJ}^{ab}(x) + U_H^{ab}(x) + U_{KP}^{ab}(x)\right]. \qquad (10)$$

The dynamics of each bead therefore obeys the Langevin equation

$$m\frac{d^2\boldsymbol{r}_a}{dt^2} = -\gamma\frac{d\boldsymbol{r}_a}{dt} - \nabla U^a(x) + \boldsymbol{\xi}^a(t) \qquad (11)$$

where γ is a friction coefficient and $\boldsymbol{\xi}^a(t)$ is stochastic delta-correlated noise which obeys the fluctuation dissipation relationship $\langle\xi_{\alpha,a}(t)\xi_{\beta,b}(t')\rangle = 2\gamma k_BT_L\delta(t-t')\delta_{\alpha\beta}\delta_{ab}$ where the Latin indexes run over particles and Greek indexes over Cartesian components.

Simulations units can be mapped to real ones by considering that the polymer beads can coarse-grain groups of nucleosomes. The nominal size of the beads can is here considered to be 30 nm, as capturing the thickness of reconstituted chromatin fibre [41]. The typical timescales over which each bead diffuses its own size is therefore

$$\tau_{Br} = \frac{\sigma^2}{D_{\text{self}}} = \frac{3\pi\eta\sigma^3}{k_BT_L} \simeq 0.01\text{s}, \qquad (12)$$

where we used the Einstein relation $D_{\text{self}} = k_BT_L/3\pi\eta\sigma$ and we considered $\eta = 150cP$ as the effective viscosity.

The recolouring process occurs at inverse rate $\tau_R = 10^3\tau_{Br}$. In each attempt, a bead is selected at random and a colour change is proposed; the move is then accepted if satisfying the Metropolis criterion as described in the text. In other words, on average, all beads will have been attempted to change their colour every $\tau_R = 10^3\tau_{Br}$ steps.

For simplicity, we model reader and writer proteins implicitly. This is effectively done by setting attractive pairwise interactions between beads bearing the same colour and by assuming a uniform selection probability during the recolouring attempts. In other words, the physical proximity of reader and writers is not required for bridging segments and changing the beads colour. This model is therefore a good approximation for the case in which reader and writer proteins saturate the solution. On the other hand, lowering the concentration of reader and writer proteins would effectively slow down the kinetics towards the equilibrium state: this situation is more complicated and we defer it to future studies.

Knotted Soliton

To test the stability of the knotted soliton we consider the time-averaged degree of epigenetic coherence $\overline{m}(x)$ at position x relative to the centre of knotted arc, $c_{\mathcal{K}}$. This can be defined as

$$\overline{m}(x) = \frac{1}{\Delta\tau}\int_{\tau_1}^{\tau_1+\Delta\tau} m(l,t)\delta(l-(c_{\mathcal{K}}+x))dt, \qquad (13)$$

where $\Delta\tau$ is the typical time between epigenetic "switching", where the magnetisation of the whole knot changes sign (such as the one occurring in Fig. **4** at about $2\ 10^5\tau_{Br}$ and $3.5\ 10^5\tau_{Br}$) and $m(l,t)$ is the magnetisation at position l and time t. In Fig. **4**(C) we report $\overline{m}(x)$ for three system replicas and compare it with the distribution of knot sizes obtained from the 3D polymer configurations (inset). As one can notice, the length of the ordered domain agrees with the length of the knotted arc. Further, the shape of the magnetisation profile $\overline{m}(x)$ displays a broad peak, which is characteristic of soliton-type solutions.

Supplementary Movies

- `M1_CO.avi` Movie showing chromatin collapsing into the compact-epigenetically ordered state corresponding to the kymograph in Fig. **2**(B).
- `M2_SD.avi` Movie showing chromatin retaining the stretched-epigenetically disordered state corresponding to the kymograph in Fig. **2**(A).
- `M3_MD.avi` Movie showing the emergence of a multi-domain configuration due to the simultaneous formation of nucleation points corresponding to the kymograph in Fig. **3**(B) and the snapshot in Fig. **3**(E).
- `M4_Knot41_1.avi` Movie showing the diffusion of a knotted soliton along stretched chromatin corresponding to the full kymograph in Fig. **4**(A).
- `M5_Knot41_2.avi` Movie showing the diffusion of a knotted soliton along stretched chromatin corresponding to a section of the kymograph in Fig. **4**(D).