bioRxiv preprint doi: https://doi.org/10.1101/058933; this version posted June 14, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

A Polymer Model with Epigenetic Recolouring Reveals a Pathway for the *de novo* Establishment and 3D Organisation of Chromatin Domains

D. Michieletto[1], E. Orlandini[2] and 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, Università di Padova, Via Marzolo 8, Padova, Italy.*

One of the most important problems in development is how epigenetic domains can be first established, and then maintained, within cells. To address this question, we propose a framework which couples 3D chromatin folding dynamics, to a "recolouring" process modelling the writing of epigenetic marks. Because many intra-chromatin interactions are mediated by bridging proteins, we consider a "two-state" model with self-attractive interactions between two epigenetic marks which are alike (either active or inactive). This model displays a first-order-like transition between a swollen, epigenetically disordered, phase, and a compact, epigenetically coherent, chromatin globule. If the self-attraction strength exceeds a threshold, the chromatin dynamics becomes glassy, and the corresponding interaction network freezes. By modifying the epigenetic read-write process according to more biologically-inspired assumptions, our polymer model with recolouring recapitulates the ultrasensitive response of epigenetic switches to perturbations, and accounts for multi-domain conformations, strikingly similar to the topologically-associating-domains observed in eukaryotic chromosomes.

INTRODUCTION

The word *"epigenetics"* refers to heritable changes in gene expression that occur without alterations of the underlying DNA sequence [1, 2]. It is by now well established that such changes often arise through biochemical modifications occurring within histone proteins – these are proteins which bind to eukaryotic DNA to form nucleosomes, the building blocks of the chromatin fibre [1]. These modifications, or "epigenetic marks", are currently thought of as forming a "histone-code" [3], which ultimately regulates expression [4].

It is clear that this histone-code has to be established *de novo* during cell development and inherited after each cell cycle through major genetic events such as replication and mitosis, or cell division [5]. A fundamental question in cell biology and biophysics is, therefore, how certain epigenetic patterns are established, and what mechanism can make them heritable. One striking example of epigenetic imprinting is the "X chromosome inactivation", which refers to the silencing of one of the two X chromosomes within the nucleus of mammalian female cells – this is crucial to avoid over-expression of the genes in the X chromosomes, which would ultimately be fatal for the cell. While the choice of which chromosome should be inactivated is stochastic within embryonic stem cells, it is faithfully inherited in differentiated cells [6]. The inactivation process is achieved, in practice, through the spreading of repressive histone modifications [7–9]. This is an example of an "epigenetic switch", a term which generically refers to the up or down-regulation of specific genes in response to, *e.g.*, seasonal changes [10–12], dietary restrictions [13], aging [14], parental imprinting [15], etc.

Although one of the current paradigm of the field is that the epigenetic landscape and 3D genome folding are intimately related [16–20], most of the existing biophysical studies incorporating epigenetic dynamics have focused on 1-dimensional or mean field models [21–28]. While these models can successfully explain some aspects of the establishment, spreading, and stability of epigenetic marks, they largely neglect the underlying 3-dimensional spatial organisation and dynamics of chromatin. This may, though, be a key aspect to consider: for instance, repressive epigenetic modifications are thought to correlate with chromatin compaction [1, 25], therefore it is clear that there must be a strong feedback between the self-regulated organisation of epigenetic marks and the 3D folding of chromatin. In light of this, here we propose a polymer model of epigenetic switches, which directly couples the 3D dynamics of chromatin folding to the 1D dynamics of epigenetics spreading.

More specifically, we start from the observation that there are enzymes which can either "read" or "write" epigenetic marks (Fig. **1**). The "readers" are multivalent proteins [17] which bridge chromatin segments with the same histone marks. The "writers" are enzymes that are responsible for the establishment and propagation of a a specific epigenetic mark, perhaps while performing facilitated diffusion along chromatin [29]. There is evidence that writers of a given mark are recruited by readers of that same mark [12, 21, 22, 24, 25, 30–32], thereby creating a positive feedback loop which can sustain epigenetic memory [22]. For example, a region which is actively transcribed by an RNA polymerase is rich in active epigenetic marks (such as the H3K4-methylated marks) [30, 33]: the polymerase in this example is "reader" which recruits the "writer" Set1/2 [33, 34]. Likewise, the *de novo* formation of centromeres in human nuclei occurs through the creation of the centromere-specific nucleosome CENP-A (a modified histone, which can thus be viewed as an "epigenetic mark") via the concerted action of the chaperone protein HJURP (the "writer") and the Mis18 complex (the "reader") [32]. Other examples of this read-write mechanism are shown in Fig. **1**. This mechanism creates a route through which epigenetic marks spread

bioRxiv preprint doi: https://doi.org/10.1101/058933; this version posted June 14, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

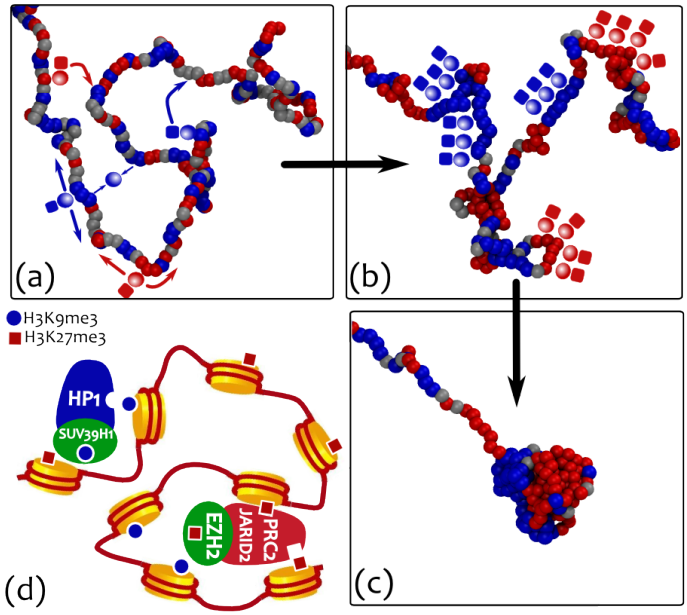

Figure 1. A 3D polymer model with "recolouring" for the propagation of epigenetic marks. (a)-(c) Multivalent binding proteins, or "readers" (transparent spheres), bind to specific histone modifications and bridge between similarly marked segments (distinguished here via their "color"). Histone-modifying enzymes, or "writers" (solid squares), are here assumed to be chaperoned by the bridge proteins. The writing (or "recolouring") activity is a consequence of 3D contiguity (perhaps through facilitated diffusion [29]) which is here modelled as a Potts-like interaction between spatially proximate monomers [35] **(a)**. The positive feedback mechanism and competition between different epigenetic marks results in a regulated spreading of the modifications **(b)** which, in turn, drives the overall folding of the polymer **(c)**. A sketch of a biological reading-writing machinery is shown in **(d)**. Heterochromatin binding protein HP1 is known to recruit methyltransferase proteins (*e.g.*, SUV39H1) which in turn trimethylate lysine 9 on histone 3 (H3K9me3) [25, 33, 36]. Similarly, the Polycomb Repressive Complex (PRC2) is known to comprise histone H3 Lys 27 (H3K27) methyltransferase enzymme EZH2 [12, 33, 37] while binding the same mark through the interaction with JARID2 [37, 38].

to spatially proximate regions along chromatin, which is responsible for the coupling between the 3D and 1D dynamics, which our model addresses.

Here we find that, for the simplest case of only 2 epigenetic states which symmetrically compete with each-other (*e.g.*, corresponding to active or inactive chromatin), our model predicts a first-order-like phase transition between a swollen, epigenetically disordered, phase, and a collapsed, epigenetically coherent, one. The first-order nature of the transition, within our model, is due to the coupling between 3D and 1D dynamics, and is important because it allows for a bistable epigenetic switch, that can retain memory of its state. When quenching the system to well below the transition point, we observe a faster 3D collapse of the model chromatin; surprisingly, we also observe a slower 1D epigenetic dynamics. We call this regime a "glassy" phase, which is characterised, in 3D, by a frozen network of strong and short-ranged intra-chain interactions – these give rise to dynamical frustration and the observed slowing down. If the change from one epigenetic mark into the other requires going through an intermediate epigenetic state, we find two main results. First, a long-lived metastable mixed state (MMS), previously absent, is now observed: this is characterised by a swollen configuration of the underlying chain where all epigenetic marks coexist. Second, we find that the MMS is remarkably sensitive to external local perturbations, while the epigenetically coherent states, once established, still display robust stability against major re-organisation events, such as replication. This behaviour is reminiscent of the features associated with epigenetic switches.

We conclude our work by looking at the case in which the epigenetic writing is an ATP-driven, and hence a non-equilibrium process. In this case, detailed balance is explicitly broken so that there can be no thermodynamic mapping of the underlying stochastic process. This case leads to a further possible regime, characterised by the formation of localised 3D compact chromatin blobs, each associated with a distinct epigenetic domain. This structure is reminiscent of the "topologically associating domains" (TAD), experimentally observed in chromosomal contact maps [39]. It is also qualitatively different from the glassy state mentioned above, because the TAD-like state displays much larger epigenetic domains in 1D.

MODELS AND METHODS

We model the chromatin fibre as a semiflexible bead-and-spring chain of M beads of size σ. To fix the ideas, we can set $\sigma = 10$ nm, approximately corresponding to a single nucleosome: in the absence of attractive interactions, this corresponds to an open chromatin fibre with no higher order structure [1]. [Note that assuming a different mapping to physical lengthscale does not affect any of the results we report below.] To each bead, we assign a "color" q representing the possible epigenetic states (marks). Here we consider $q \in \{1, 2, 3\}$, *i.e.* three possible epigenetic marks such as methylated (inactive), unmarked (intermediate) and acetylated (active).

In addition to the standard effective potentials to ensure chain connectivity (through a harmonic potential between consecutive beads) and bending rigidity (through a Kratky-Porod potential [41]), we consider a repulsive/attractive interactions mediated by the epigenetic marks (colors). This is described by a truncated-and-shifted Lennard-Jones potential, defined as follows,

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

whereas $U_{LJ}^{ab}(x) = 0$ for $x > x_c^{q_a q_b}$. In Eq. (10), $\mathcal{N}$ is a normalisation constant and the parameter ϵ_{ab} is set so that $\epsilon_{ab} = \epsilon$ for $q_a = q_b$ and $\epsilon_{ab} = k_B T_L$ otherwise. The q-dependent interaction cut-off $x_c^{q_a q_b}$ is given by $2^{1/6}\sigma$, to

bioRxiv preprint doi: https://doi.org/10.1101/058933; this version posted June 14, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

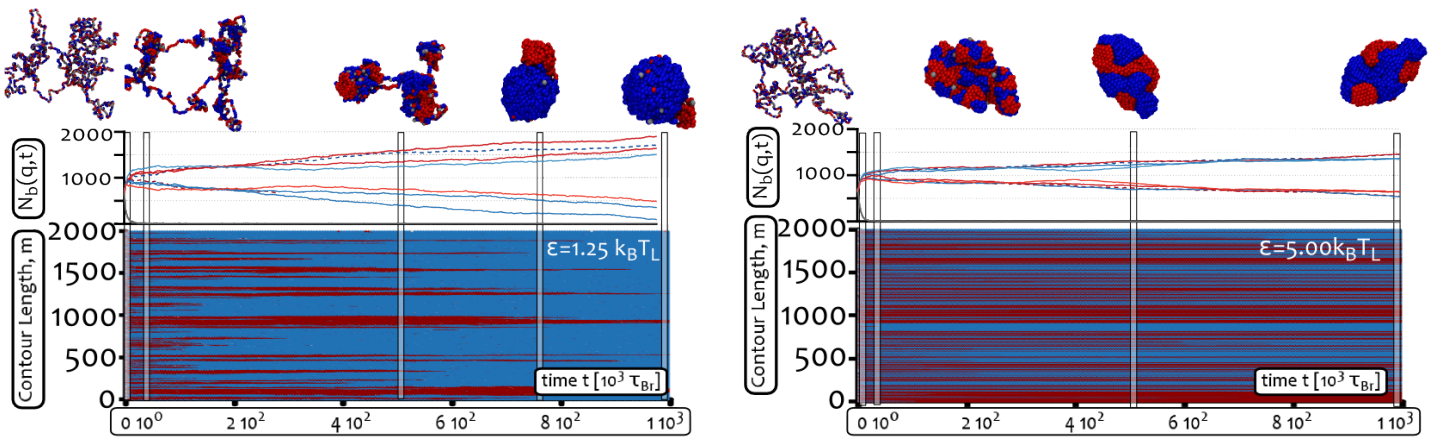

Figure 2. **The two-state model displays a symmetry-breaking between epigenetic coherent states.** Top row: snapshots of some of the 3D configurations assumed by the polymers as a function of time, and for two choices of $\alpha = \epsilon/k_BT_L$. Middle row: time evolution of the total number of beads of type q, $N_b(q,t)$, for four independent trajectories (the dashed one corresponds to the trajectory from which the snapshots are taken). Bottom row: time evolution of the color of each polymer bead, as a "kymograph" [40]. The simulations show a symmetry-breaking phenomenon, in which one of the two interacting states (either red or blue) takes over the whole polymer. Higher values of the interaction parameter lead to a faster 3D collapse followed by a slower epigenetic transition from the disorder (multi-phase) to ordered (single-phase) state (see also Suppl. Movies M1-M3). Contact maps corresponding to each snapshot are reported in SI Figs. S2.

model steric repulsion, or $R_i > 2^{1/6}\sigma$ to model attraction. [Here, we consider $R_i = 1.8\sigma$, which simultaneously ensures short-range interaction and computational efficiency.] In what follows, the cut-offs are chosen so that beads with different colors, or with color corresponding to no epigenetic marks (i.e., $q = 3$), interact via steric repulsion, whereas beads with the same color, and corresponding to a given epigenetic mark (*e.g.*, $q = 1$, or $q = 2$), self-attract, modelling interactions mediated by a bridging protein, one of the "readers" [1].

The time evolution of the system is obtained by coupling a 3D Brownian polymer dynamics at temperature T_L, with a recolouring Monte-Carlo dynamics of the beads which does not conserve the number of monomer types. recolouring moves are proposed every $\tau_{\rm Rec} = 10^3\tau_{Br}$, where τ_{Br} is the Brownian time associated with the dynamics of a single polymer bead: they are performed by attempting M changes of the beads colors. Each color change is accepted according to the standard Metropolis acceptance ratio with effective temperature $T_{\rm Rec}$ and Potts-like energy difference computed between beads that are spatially proximate (*i.e.*, within distance R_i in 3D). It is important to notice that, whenever $T_L \neq T_{\rm Rec}$, detailed balance of the full dynamics is broken, which may be appropriate if epigenetic spreading and writing depend on non-thermal processes (*e.g.*, if they are ATP-driven). More details on the model, and values of all simulation parameters, are given in the SI.

The model we use therefore couples the Ising-like (or Potts-like) epigenetic recolouring dynamics, to the 3-dimensional kinetics of polymer folding (or of polymer collapse). In most simulations we consider, for simplicity, $T_L = T_{\rm Rec}$, and we start from an equilibrated chain configuration in the swollen phase (*i.e.*, at very large T_L), where beads are randomly colored with uniform probability. The polymer and epigenetic dynamics is then studied after reducing T_L to values below the polymer collapse temperature (Θ temperature). In practice, in what follows, we set $\epsilon/(k_BT_L) \geq 1$ unless otherwise stated.

RESULTS

An effective two-state model leads to a first-order-like transition, and epigenetic bistability

For simplicity, we focus here on the case in which three states are present, but only two of them ($q = 1$, red and $q = 2$, blue) are self-attractive, while the third is a neutral state that does not self-attract, but can participate to coloring dynamics ($q = 3$, grey). Transition between any two of these three states are possible in this model. Because we find that the grey (unmarked) state loses out to the self-attractive ones, we refer to this as an effectively "two-state" model. This scenario represents the case with two competing epigenetic marks (e.g., an active acetylation mark and an inactive methylation mark), while the third state stands for unmarked chromatin.

Fig. **2** reports the joint polymer and epigenetic dynamics (starting from the swollen randomly colored initial state), for two different values of $\alpha = \epsilon/k_BT_L$. The global epigenetic recolouring is described via $N_b(q,t)$, the total number of beads in a certain state q at time t; the local epigenetic dynamics is instead represented by a "kymograph" [40] (Fig. **2**).

It is readily seen that the chain condenses fairly quickly into a single globule and clusters of colors emerge and coarsen. Differently-colored clusters then compete, so that ultimately we are left with an epigenetically coherent globular phase. Because the red-red and blue-blue interactions are equal, the selection of which epigenetic mark dominates is via symmetry-breaking of the red$\leftrightarrow$blue ($\mathbb{Z}_2$) sym-

bioRxiv preprint doi: https://doi.org/10.1101/058933; this version posted June 14, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

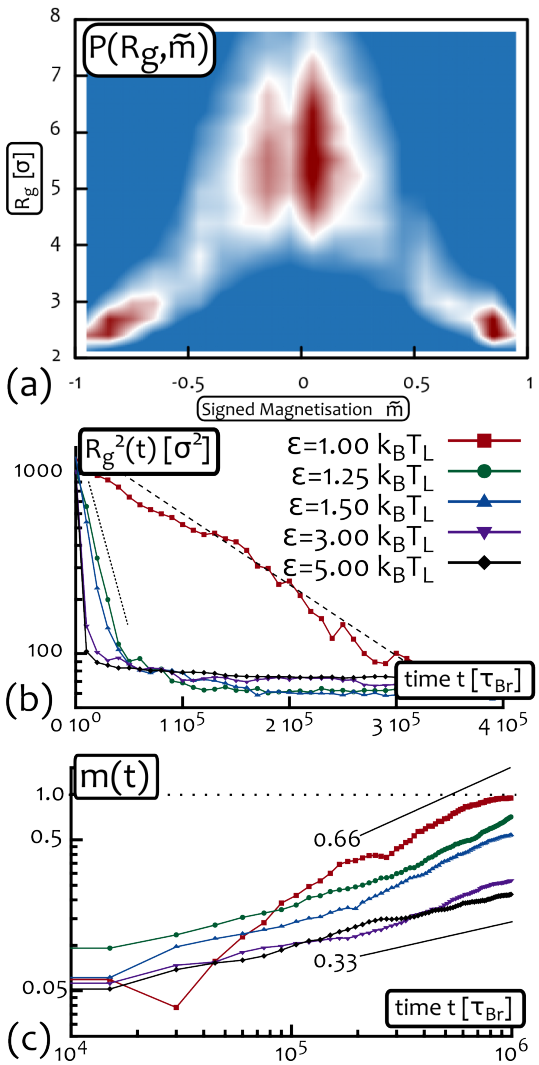

Figure 3. The "two-state" model displays a discontinuous transition driven by an α-dependent epigenetic dynamics. **(a)** Plot of the joint probability $P(R_g, \tilde{m})$ for a chain of $M = 50$ beads, obtained from 100 independent simulations of duration $2\ 10^6 \tau_{Br}$. **(b)** Time evolution of the gyration radius R_g of the polymer. **(c)** Plot of the epigenetic magnetisation (defined in eq. (3)) as a function of time. One can readily appreciate two main features of this model. First, at the critical point $\alpha_c = \epsilon/k_B T_L|_c$ (roughly at $\alpha_c = 1.15$ for $M = 50$), the system displays a discontinuous phase transition. The two phases (disordered, low-$\tilde{m}$ and large R_g and ordered, high-$\tilde{m}$ and small R_g) coexist, *i.e.* the joint probability displays three maxima (equivalently, the effective free energy $F = -k_B T_L \log P(R_g, \tilde{m})$ displays three minima). Second, larger values of α lead to a faster polymer collapse dynamics (faster decay of R_g), combined to a slower recolouring dynamics towards the epigenetically coherent state (slower growth of $m(t)$).

metry. The transition between the swollen-disordered and collapsed-coherent phases bears the hallmark of a discontinuous, first-order-like transition: for instance, we observe coexistence of the two phases, as well as some marked hysteresis. To pinpoint the transition, we calculate the joint probability $P(R_g, \tilde{m})$ of observing a state with a given value of gyration radius, R_g, and signed "epigenetic magnetisation",

$$\tilde{m} \equiv \frac{1}{M}\left(N_b(q=1) - N_b(q=2)\right). \qquad (2)$$

When $P(R_g, \tilde{m})$ is computed at large times after the quench (see Fig. **3** and SI, Fig. S3), the single maximum expected for the swollen-disordered phase (large R_g and small $\tilde{m}$) splits, below the transition point $\alpha_c = \epsilon/k_B T_L|_c$, into two symmetric maxima corresponding to the collapsed-ordered phase (small R_g and $\tilde{m} \simeq \pm 1$). More importantly, at the transition three maxima are clearly visible suggesting the presence of phase coexistence (see Fig. **3** and SI Figs. S3-S4).

The existence of a first-order-like transition with hysteresis is important, since it naturally provides a framework within which epigenetic states can be established and maintained in the presence of external fluctuations. It also provides a marked difference between our model and previous ones, which approximated the epigenetic (recolouring) dynamics as a one-dimensional process, where nucleosome recruitment was regulated by choosing an *ad hoc* long-range interaction [21, 28]. These effectively 1D models display either a second order transition [21, 42, 43], or a first-order transition, but only in the mean-field ("all against all") [28]. In our model the first-order-nature of the transition critically requires the coupling between the 3D polymer collapse and the 1D epigenetic dynamics – in this sense, the underlying physics is similar to that of magnetic polymers [44].

The dynamical feedback between chromatin folding and epigenetic recolouring can be appreciated by looking at Suppl. Movies M1-M2, where it can be seen that local epigenetic fluctuations trigger local chromatin compaction. Suppl. Movies M1-M2 also show that the dynamics of the transition from swollen to globular phase is, to some extent, similar to that experienced by a homopolymer in poor solvent conditions [45–52]. namely a formation of small compact clusters along the chain (pearls) that eventually coalesce into a single globule. Unlike the homopolymer case, however, the pearls may be differently colored giving rise at intermediate or late times to frustrated dynamics, where two or more globules of different colors compete through strong surface tension effects. When several globules are present, we observe cases in which two or more pearls of the same color, that are distant along the chain but close in 3D, merge by forming long-ranged loops (see snapshots in Fig. **2**, contact maps in SI and Suppl. Movies M1-M2).

Deep quenches into the collapsed phase lead to a slow, glassy, dynamics associated with a frozen, short-ranged interaction network

An intriguing feature observed in the collapse-recolouring dynamics is that quenching at different temperatures affects non trivially the timescales of chromatin condensation and epigenetic evolution towards a single coherent state (see also Suppl. Movie M3. The separation between these two timescales increases with α (*i.e.*, for deeper quenches), as can be readily seen in Fig **3**, where we compare the time evolution of the mean squared radius of gyration of the chain

bioRxiv preprint doi: https://doi.org/10.1101/058933; this version posted June 14, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

$R_g^2(t)$ and the time-dependent (absolute) epigenetic magnetisation

$$m(t) = \frac{1}{M} \left| N_b(q=1,t) - N_b(q=2,t) \right| , \qquad (3)$$

for different values of α.

While R_g decays exponentially with a timescale that decreases as α increases (Fig. **3**(a)), the epigenetic magnetisation grows as $m(t) \sim t^\beta$, where the dynamical exponent β decreases from $\simeq 2/3$ to $\simeq 1/3$ as α increases. Note that the value 2/3 has been reported in the literature as the one characterising the coarsening of pearls in the dynamics of homopolymer collapse [47]. The fact that in our model this exponent is obtained for low values of α suggests that in this regime the timescales of polymer collapse and epigenetic coarsening are similar. In this case, we expect $m(t)$ to scale with the size of the largest pearl in the polymer, whose color is the most likely to be selected for the final domain – *i.e.*, the dynamics is essentially determined by the homopolymer case. Our data are instead consistent with an apparent exponent smaller than 2/3 for larger α, signalling a slower epigenetic dynamics.

The interesting finding that a fast collapse transition gives rise to a slowing down of the recolouring dynamics can be understood in terms of the evolution of the network of intra-chain contacts. This can be monitored by defining the interaction matrix

$$P_{ab}(t) = \begin{cases} 1, \text{ if } d_{ab}(t) < R_i \\ 0, \text{ otherwise} \end{cases}$$

where $a, b = 1, \ldots, M$ denote two monomers, and $d_{ab}(t) = |\boldsymbol{r}_a(t) - \boldsymbol{r}_b(t)|$. From the interaction matrix we can readily obtain useful informations on the network structure, such as the average number of neighbours per bead,

$$N_n(t) = \frac{1}{M} \sum_{a \neq b} P_{ab}(t) \qquad (4)$$

or the average "spanning distance", which quantifies whether the network is short- or long-ranged (see SI for details). The contact probability between beads a and b can also be simply computed, as the time average of $P_{ab}(t)$.

As expected, for larger values of α, $N_n(t)$ saturates to a maximum value. On the other hand, and more importantly, for higher values of the interaction strength α, a dramatic change in the spanning distance is observed. This effect is well captured by plotting a network representation of the monomer-monomer contacts, as reported in Fig. **4** (see SI, Figs. S6-S9 for a more quantitative analysis). This figure shows that at large α there is a depletion of the number of edges connecting distant monomers along the chain, while short-ranged contacts are enhanced (see caption of Fig. **4** for details; see also contact maps in SI Fig. S2). Note that this finding is consistent with the fractal globule conjecture [53], for which a globule obtained by a fast collapse dynamics is rich of local contacts and poor in non-local ones. However, the present system represents a novel instance of "annealed" collapsing globule, whose segments are dynamically recoloured as it folds.

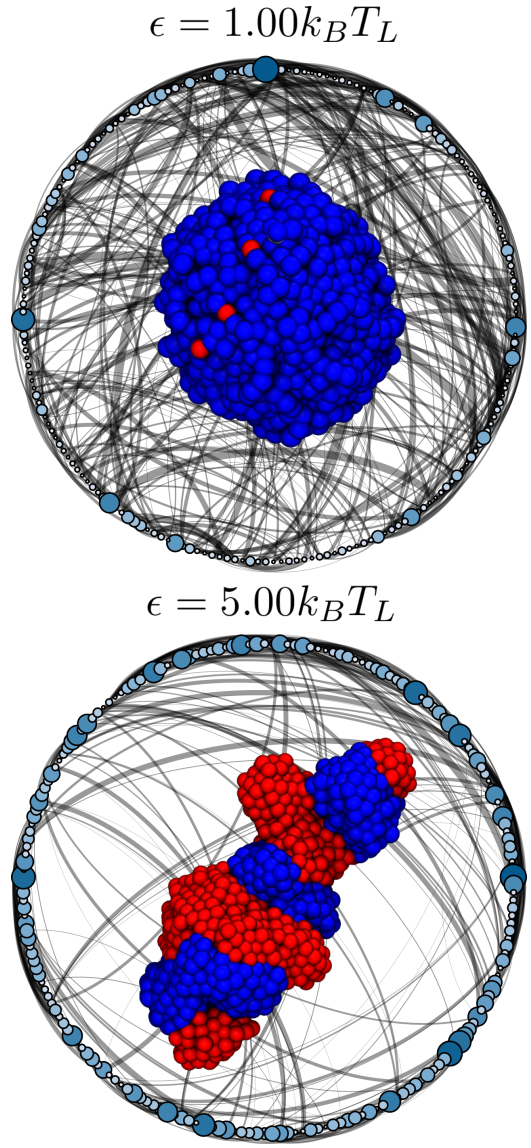

Figure 4. The network of interactions is short ranged for fast collapsing coils. Snapshot of the network of bead-bead contacts taken at $t = 10^6 \tau_{Br}$ for two simulations with (top) $\epsilon = 1 k_B T_L$ and (bottom) $\epsilon = 5 k_B T_L$. For sake of clarity of visualisation, each node of the network coarse grains 10 beads along the chain. Node size and color intensity encodes the number of interactions within the coarse-grained monomers. Edges are only drawn between nodes which contain interacting monomers, and their thickness is proportional to the (normalised) number of contacts. To improve the visualisation, only edges corresponding to a contact probabilities between monomers (see text) in the top 30% are displayed. Snapshots of the respective 3D conformations are also shown. It is important to notice that higher values of α lead to short-ranged networks, *i.e.* fewer edges but larger nodes.

Finally, in order to characterise the change in the kinetics of the network, we quantify the "mobility" of the contacts, or the "neighbour exchange rate", following polymer collapse. We therefore compute

$$\kappa_n(t) = \frac{1}{M} \sum_{a \neq b} \left[1 - \delta(P_{ab}(t), P_{ab}(t - \Delta t)) \right] , \qquad (5)$$

bioRxiv preprint doi: https://doi.org/10.1101/058933; this version posted June 14, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

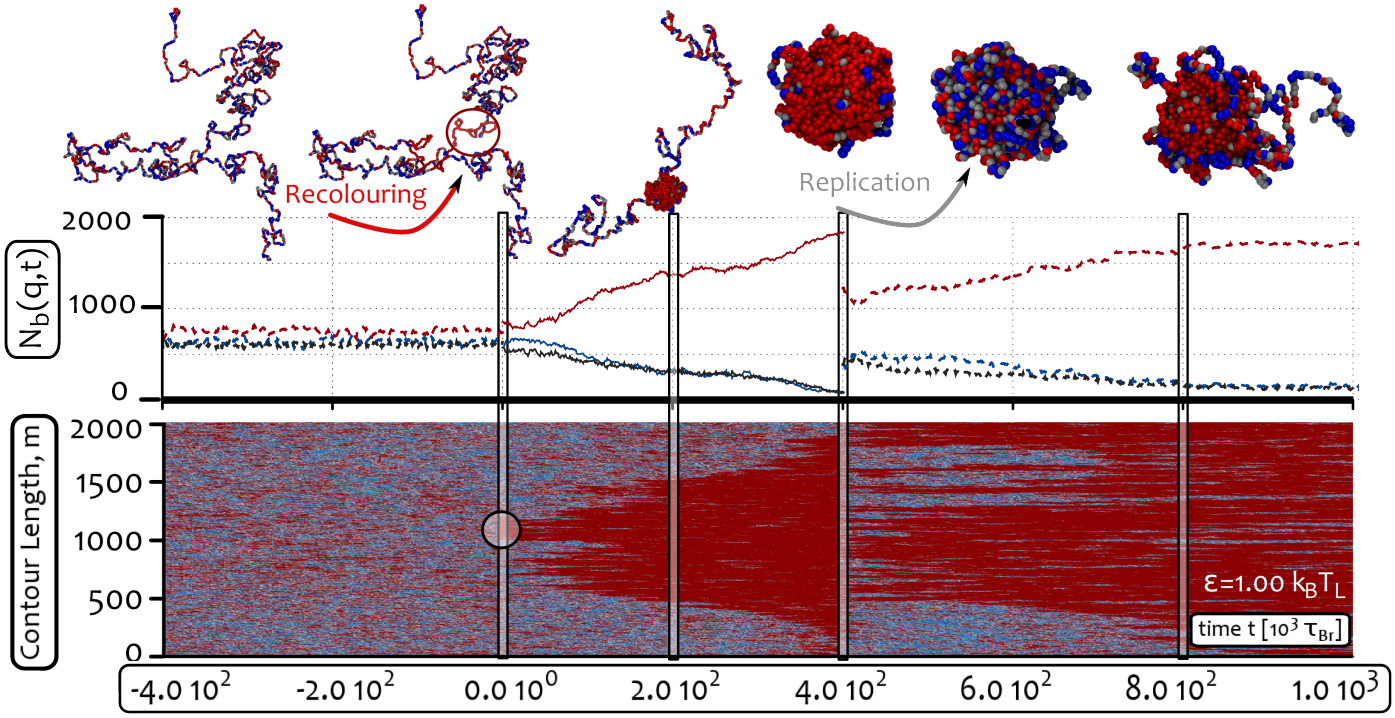

Figure 5. The "two-state with intermediate state" model displays ultra-sensitivity. Time evolution of the system starting from a mixed metastable state (MMS): At $t = 0$ a localised perturbation of the MMS is imposed artificially by recolouring a segment of 200 beads (10% of polymer length). This perturbation gives rise to a transition towards the collapsed and epigenetically coherent state. At time $t = 4\ 10^5\ \tau_{Br}$ we then performed a semi-conservative replication on the collapsed phase. This is achieved by assigning a random color to 50% of the beads all along the polymer. Following this extensive (*i.e.* non local) colour perturbation, one can observe the polymer going back to the ordered phase; this suggests a broad stability of the epigenetically coherent state, in contrast to the more sensitive MMS. See also Suppl. Movie M4, where we show the whole dynamics. Contact maps are shown in SI Fig. S11.

where $\Delta t = 10^3 \tau_{Br} = \tau_{\rm Rec}$ is the gap between two measurements. We find that above $\alpha = 3$, the time-averaged value of the neighbour exchange rate, normalised by the average number of neighbours, $\langle \kappa_n \rangle / \langle N_n \rangle$, sharply drops from values near unity, indicative of mobile rearranging networks, to values close to zero, signalling a frozen network or contacts (see SI Fig. S10).

The "topological freezing" (see also Suppl. Movie M3) due to fast folding is also partially reflected by the strongly aspherical shapes taken by the collapsed coils in the large α regime (see snapshots in Fig. **2** and Fig. **4**).

The emerging scenario is therefore markedly different from the one suggested in models for epigenetic dynamics with long-range [21, 42, 43] or mean-field interactions [28], where any two beads in the chain would have a finite interaction probability. Instead, in our case, this is only a valid approximation at small α, whereas at large α a given bead interacts with only a subset of other beads (see Fig. S6), and it is only by averaging over different trajectories and beads that we get the power-law decay of the contact probability assumed in those studies (see Fig. S7). This observation is, once again, intimately related to the fact that we are explicitly taking into account the 3D folding together with the epigenetic dynamics.

An intermediate state in the epigenetic dynamics leads to long-lived mixed metastable states and ultrasensitive epigenetic switches

Up until now, our model has been based on a simple rule for the epigenetic dynamics, where each state can be transformed into any other state. It is of interest to extend the model by introducing a specific chemical chemical pathway for the epigenetic writing, as previously done in 1D models [26]. A typical example in epigenetics is when a nucleosome, with a specific epigenetic mark (corresponding to, say, the blue "state"), can be converted into another state (say, the "red" one) only after the first mark has been removed. This two-step re-writing mechanism can be described by considering a "neutral" or "intermediate" state (IS) through which any nucleosome has to transit before changing its epigenetic state (say, from "blue" to "red") [21, 26]. Previous studies, based on mean field or *ad hoc* power law interaction rules for the recruitment of epigenetic marks have shown that the presence of such an intemediate unmakred state can enhance bistability and create a long-lived mixed metastable state (MMS), in which all epigenetic states coexist in the same system [21, 23, 26].

Differently from the simulations reported in the previous Sections, where we never observed a long-lived mixed state, as the "red" or "blue" beads rapidly took over the "grey" beads, in this case we observe that the mixed state is

bioRxiv preprint doi: https://doi.org/10.1101/058933; this version posted June 14, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

metastable. This finding is in agreement with the 1D models of Refs. [21, 26], although in our case the cooperativity is a natural consequence of the folding dynamics.

In particular, we find that a mixed metastable state (MMS) is observed with 40% of probability at $\alpha = 1$ whereas for higher values of α this state is unstable and never observed. A typical example of a MMS is reported in the early times of Fig. **5**: one can see that it is characterised by a swollen coil with no sign of epigenetic domains, and all three states coexist in the same configuration.

In order to study the stability of the MMS, we perturb the system by artificially recolouring (in a coherent fashion) a localised fraction (10%) of beads along the chain. From Fig. **5** one can see that, after the perturbation (performed at $t = 0$), the chain forms a nucleation site around the artificially recoloured region that eventually grows as an epigenetically coherent globule. The spreading of the local epigenetic domain throughout the whole chain can be followed from the kymograph in Fig. **5** (see also Suppl. Movie M4 and contact maps in SI, Fig. S11).

Note that, in practice, for an *in vivo* chromatin fibre, this local coherent recolouring perturbation might be due, *e.g.*, to an increase in local concentration of a given "writer" (or of a reader-writer pair): our result therefore show that a localised increase could be enough to trigger an extensive epigenetic response, or switch, which might affect a large chromatin region.

To test the stability of the coherent globular state following the symmetry breaking, we then performed a random recolouring of 50% of the beads (*i.e.*, 50% of the beads are randomly given one of the three possible colors). This perturbation is chosen because it mimics qualitatively how epigenetic marks may be semi-conservatively passed on during DNA replication.

After this instantaneous extensive and random recolouring (performed at $t = 4\ 10^5\ \tau_{Br}$ in Fig. **5**), we observe that the model chromatin returned to the same ordered state, suggesting that the epigenetically coherent state, once selected, is robust to even extensive recolouring perturbations (see also Suppl. Movie M4).

In the full simulation reported in Figure **5**, we therefore first perform a small perturbation on a MMS, which drives the system towards a globular, epigenetically coherent state; following this, we partially randomize the epigenetic marks, finding that this does not destabilise the ordered state. This largely asymmetric response of the system to external fluctuations, which depends on its instantaneous state, is known as "ultra-sensitivity" [22].

As we discuss more in the Conclusions, ultrasensitivity is a highly desirable feature in epigenetic switches and during development. A striking example of this feature is the previously mentioned X-chromosome inactivation in mammalian female embryonic stem cells. While the selection of the chromosome copy to inactivate is stochastic at the embryonic stage, the choice is then epigenetically inherited in committed daughter cells [6]. Thus, one may imagine that a small perturbation may turn the whole chromosome from a mixed metastable state into an inactive heterochromatic state (*e.g.*, an "all-red" state in terms on Fig. **5**); when chromatin replicates, this can be viewed as an extensive epigenetic fluctuation, which, over time, decays to leave the same "red" heterochromatic stable state.

From a physics perspective, the results reported in this section and encapsulated in Figure **5** are of interest because they show that the presence of the intermediate state, while not affecting the robustness of the steady states, dramatically changes the kinetics of the system by introducing the additional mixed metastable state [26].

Breaking detailed balance can stabilise a multi-pearl, TAD-like, chromatin structure in steady state

In the previous Section we have considered the case in which 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 density. This is certainly a simplification because the epigenetic writing is in general a non-thermal, out-of-equilibrium process, which entails biochemical enzymatic reactions with chromatin remodelling and ATP consuption [1]. Thus, it is important to see what is the impact of breaking detailed balance in the dynamics of our model.

We address this point by considering a recolouring temperature $T_{\rm Rec}$ that differs from the polymer dynamic temperature T_L. In what follows we set $T_{\rm Rec} = 0.1\epsilon/k_B$, keeping $T_L = 1.75\epsilon/k_B$ (*i.e.*, $\alpha = \epsilon/k_B T_L = 0.57$, see SI for other cases). Since $T_{\rm Rec} \neq T_L$, one can readily show, through the Kolmogorov criterion, that the detailed balance is violated by computing the net probability flow along a closed loop through some states of the system (see SI).

By considering a polymer temperature above the Θ point of an homopolymer, and a recolouring effective temperature well below T_L, we can probe a regime in which the polymer is swollen, yet the epigenetic interactions are large enough to induce locally coherent states. This leads to a competition between polymer entropy, which favours the disordered phase, and 1D epigenetic coherency – as we shall see, such a competition generates new interesting physical behaviour. Quite remarkably, for this parameter choice, the system is not observed to collapse into a single globule, but instead folds into a multi-pearl state which is stable, or at least very long-lived (see Fig. **6**).

A simple argument to understand why an overall epigenetically coherent phase is hard to reach within a swollen phase is the following: a swollen self-avoiding walk is characterised by an intra-chain contact probability scaling as

$$P_c(m) \sim m^{-c} \tag{6}$$

with $c = (d + \theta)\nu > 2$ [54, 55]. This value implies that the interactions are too local to trigger a phase transition in the epigenetic state, at least within the Ising-like models considered in Ref. [42].

An important lengthscale characterising order in such systems is the magnetic correlation length, which quantifies the size of the epigenetic domains. We observe that this lengthscale, $\xi(T_{\rm Rec})$ – defined through the exponential decay of the

bioRxiv preprint doi: https://doi.org/10.1101/058933; this version posted June 14, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

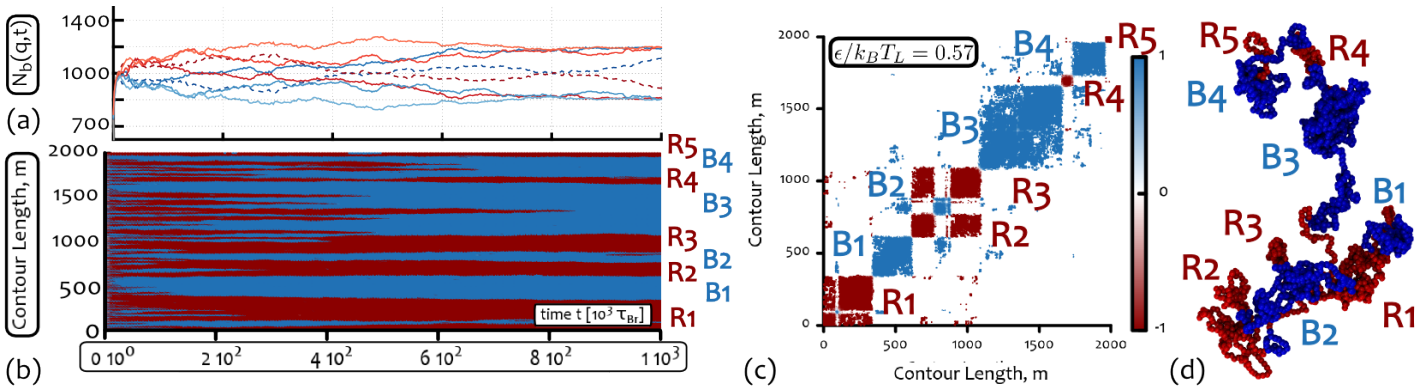

Figure 6. **Breaking Detailed Balance leads to the formation of TAD-like structures**. The figure shows: **(a)** the number of red (and blue) coloured beads $N_b(q,t)$ for four independent system replicas; **(b)** the kymograph; **(c)** a time-averaged contact map; **(d)** a snapshot of a 3D configuration in steady state. Note that, for the particular choice of α made here, the system does not show symmetry breaking, as $N_b(q,t)$ is not diverging at large times, and remains extensive for both red and blue beads. On the other hand, the system still displays folding into local domains (whose size is a significant portion of the whole chain), while maintaining an overall swollen conformation. The snapshot shown corresponds to a time equal to 10^6 τ_{Br}; the visible TAD-like structures in the snapshot and in the contact map are enumerated as in the kymograph, to ease comparison. The contact map is obtained by averaging over 40 configurations taken at different time-steps in the last part of the trajectory ($t > 8\ 10^5\ \tau_{Br}$) ; contacts between blue beads are scored as 1, while contacts between red monomers as -1, and mixed contacts as 0. See SI, Fig. S12 for instantaneous contact maps.

epigenetic correlation function (see SI) – increases as $T_{\rm Rec}$ decreases, but never spans the whole polymer length.

To understand why the multi-pearl state of Figure **6** is stable, we first notice that each epigenetic domain with typical size $\sim \xi(T_{\rm Rec})$ can be seen as a short homopolymer whose Θ temperature $T_L^{*,h}$, is higher than the one for a random copolymer of the same length; therefore, as the segments become coherently coloured, they can transiently form globules if the polymer temperature is close enough to the Θ temperature $T^{*,h}(m = \xi(T_{\rm Rec})_L)$ for a homopolymer of size $\xi(T_{\rm Rec})$.

The second element worth highlighting is that for a homopolymer m beads long, there exist a temperature range $\Delta T/T_L^{*,h} \sim m^{-1/2}$, for which the coil and globule coexist [45]. Consequently, the polymer can also be seen as a collection of de Gennes' blobs with typical size [45]

$$m^* \sim \left[(T_L - T_L^{*,h})/T_L^{*,h}\right]^{-2}. \tag{7}$$

This finally implies that in this regime there are two competing length-scales: (i) ξ tuned by $T_{\rm Rec}$ and (ii) m^* tuned by the polymer temperature T_L.

The most interesting regime, an example of which is reported in Fig.**6** ($T_L = 1.75\epsilon/k_B$ and $T_{\rm Rec} = 0.1\epsilon/k_B$), is the one where

$$\xi(T_{\rm Rec}) \lesssim m^*(T_L). \tag{8}$$

In this regime, the coherent segments can intermittently stick together to form the multi-pearl structure we observe (see Suppl. Movie M5 for a full dynamics); the typical pearl size is set by $\xi(T_{\rm Rec})$. The transient local clustering creates average 3D contiguity picked up by the average contact map in Figure **6** (instantaneous contact maps, which show that the aggregates are short-lived, are shown in SI Fig. S12).

Another possible regime (reported in the SI, Fig. S15) is the one where

$$\xi(T_{\rm Rec}) > m^*(T_L). \tag{9}$$

In this case the polymer collapses, in analogy with what we have found in the previous Sections. Finally, in the limit $\xi \to 0 \ll m^*$, the polymer returns to the disordered phase (see SI Fig. S15 and Suppl. Movies M5-M6).

In practice, by using the estimate $T_L^{*,h} \simeq 1.85\epsilon/k_B$ (see SI Fig. S13 and Fig. S14), one can see that $200 \sim \xi(T_{\rm Rec} = 0.1\epsilon/k_B) > m^*(T_L = 1.5\epsilon/k_B) \sim (1.8/0.3)^2 \simeq 25$ monomers. This regime leads to a fully collapsed state, since m^* is much smaller than ξ (SI Fig. S15). On the other hand, the case for which $200 \sim \xi(T_{\rm Rec} = 0.1\epsilon/k_B) < m^*(T_L = 1.75\epsilon/k_B) \sim (1.8/0.05)^2 \simeq 350$ beads leads to very stable and locally coherent organisation, as shown in Fig. **6**.

A final remark is in order here. The multi-pearl state we report in this work is strikingly reminiscent of the topologically-associated domains, or TADs, found in Hi-C experiments with mammalian chromosomes, and which lead to a "block-like" appearance of the contact map, not unlike that in Figure **6** [17, 56, 57]. Notwithstanding this intriguing similarity, we note that the domains that we observe arise stochastically, so that, unlike TADs, they would appear in different places along the chromatin fibre in different simulations (corresponding to different cells in a Hi-C experiment). It is also likely that the establishment of TADs in practice requires insulator elements such as CTCF, or other architectural proteins [1] which we are not considering in the simulations reported here.

Our results though are of relevant to the biophysics of TADs as they point out that understanding the mechanism through which they are established may require going beyond an effectively thermodynamic framework – within our model, a multi-pearl TAD-like state is in fact only stable

bioRxiv preprint doi: https://doi.org/10.1101/058933; this version posted June 14, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

when considering a truly non-equilibrium dynamics violating detailed balance.

DISCUSSION AND CONCLUSIONS

In summary, here we have studied a 3D polymer model with epigenetic "recolouring", which explicitly takes into account the coupling between the 3D folding dynamics of a semi-flexible chromatin fibre and the 1D "epigenetic" spreading. Supported by several experimental findings and well-established models [1, 17], we assume self-attractive interactions between chromatin segments bearing the same epigenetic mark, but not between unmarked or differently-marked segments. We also assume a positive feedback between "readers" (binding proteins aiding the folding) and "writers" (histone-modifying enzymes performing the recolouring), which is supported by experimental findings and 1D models [21, 22, 25, 33, 38, 58].

Our work couples for the first time 3D chromatin and 1D epigenetic dynamics. This synergy is crucial to consider, as it is very likely that the timescales for chromatin folding and epigenetic recolouring are comparable. The formation of local TADs in a cell in fact requires at least several minutes, while the establishment of higher order, non-local contacts, is even slower [59]; at the same time, histone-modifications occur through enzymatic reactions whose rate is of the order of at least inverse seconds [22]. A practial example could be, for instance, active epigenetic marks deposited by a travelling polymerase during the ~ 10 minutes over which it transcribes an average human gene of 10 kbp [60].

Furthermore, there are examples of biological phenomena *in vivo* which point to the importance of the feedback between 3D chromatin and quasi-1D epigenetic dynamics. A clear example is the inactivation of an active and "open" [1] chromatin region which is turned into heterochromatin. In this case, the associated methylation marks favour chromatin self-attractive interactions [60] and these in turn drive the formation of a condensed structure [1, 33] whose inner core might be difficult to reach by other freely diffusing reactivating enzymes.

Rather fittingly, we highlight that one of our main results is that the coupling between conformational and epigenetic dynamics can naturally drive the transition between a swollen and epigenetically disordered phase at high temperatures and a compact and epigenetically coherent phase at low temperatures (Fig. **2**), and that this transition is discontinuous, or first-order-like, in nature (Fig. **3**).

It is known that while purely short-range interactions cannot drive the system into a phase transition, effective (or *ad hoc*) long-range interactions within an Ising-like framework can induce a (continuous) phase transition in the thermodynamic limit [42, 43]. In our case, importantly, the transition is discontinuous (see Fig. **3** and SI, Fig. S3), and this is intrinsically related to the coupling between 3D and 1D dynamics. The physics leading to a first-order-like transition is therefore reminiscent of that at work for magnetic polymers [35] and hence fundamentally different with respect to previous work, which neglected the conformation-epigenetics positive feedback coupling.

It is interesting to notice that the discontinuous nature of the transition observed in this model can naturally account for bistability and hysteresis, which are both properties normally associated with epigenetic switches.

We note that the model reported here also displays a richness of physical behaviours. For instance, we intriguingly find that by increasing the strength of self-attraction the progress towards the final globular and epigenetically coherent phase is much slower; we characterise this glass-like dynamics by analysing the network of contacts and identifying a dramatic slowing down in the exchange of neighbours alongside a depletion of non-local contacts (see Figs. **4**). We argue that the physics underlying the emergence of a frozen network of intra-chain interactions might be reminiscent of the physics of spin glasses with quenched disorder [61] (see Figs. **4** and SI Fig. S10).

We have also shown that, by forcing the passage through an intermediate (neutral, *i.e.*, unmarked) state during the epigenetic writing, it does not change either the nature of the transition or the long-time behaviour, but it produces major effects on the dynamics. Most notably, it allows for the existence of a long-lived metastable mixed state (MMS) in which all three epigenetic states coexist. This case is interesting as it displays ultrasensitivity to external perturbations, *i.e.*, the MMS is sensitive to small local fluctuations which drive large conformational and global changes, while the epigenetically coherent states show stability against major and extensive re-organisation events such as semi-conservative chromatin replication (Fig. **5**).

Like hysteresis and bistability, ultrasensitivity is important in *in vivo* situations, in order to enable regulation of gene expression and ensure heritability of epigenetic marks in development. For instance, it is often that case that, during development a localised external stimulus (*e.g.*, changes in the concentration of a transcription factor or a morphogen) is enough to trigger commitment of a group of cells to develop into a cell type characterising a certain tissue rather than another [1]. On the other hand, once differentiated, such cells need to display stability against intrinsic or extrinsic noise. Ultrasensitivity similar to the one we observe would enable both types of responses, depending on the current chromatin state.

A further captivating example of ultrasensitive response is the previously mentioned case of the X-chromosome inactivation. Also in that case, the selection of which of the two X-chromosomes to silence is stochastic in female mammalian embryonic stem cells: specifically, it is suggested that a localised increase in the level of some RNA transcripts (XistRNA) can trigger heterochromatization of the whole chromosome (which is at that point called a Barr body) by propagating repressive marks, such as H3K27me3, through recruitment of the polycomb complex PRC2 [9]. Once the inactive X copy is selected, the choice is epigenetically inherited in daughter cells [6].

Finally, we have studied the case in which the epigenetic dynamics is subject to a different stochastic noise, with re-

bioRxiv preprint doi: https://doi.org/10.1101/058933; this version posted June 14, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

spect to the 3D chromatin dynamics. This effectively "non-equilibrium" case, where detailed balance of the underlying dynamics is broken, leads to interesting and unique physical behaviours. Possible the most pertinent is that we observe, and justify, the existence of a parameter range for which a multi-pearl state consisting of several globular domains coexist, at least for a time corresponding to our longest simulation timescales (Fig. **6**). This multi-pearl structure is qualitatively reminiscent of the topologically associated domains in which a chromosome folds *in vivo*, and requires efficient epigenetic spreading in 1D, together with vicinity to the theta point for homopolymer collapse in 3D.

Although one of the current paradigm of chromosome biology and biophysics is that the epigenetic landscape directs 3D genome folding [16–20], an outstanding question is how the epigenetic landscape is established in the first place – and how this can be reset *de novo* after each cell division. Within this respect, our results suggest that the inherent non-equilibrium (*i.e.*, ATP-driven) nature of the epigenetic read-write mechanism, can provide a pathway to enlarge the possible breadth of epigenetic patterns which can be established stochastically, with respect to thermodynamic models.

Overall, the model presented in this work can be thought of as a general paradigm to study 3D chromatin dynamics coupled to an epigenetic read-write dynamics in chromosomes. All our findings further strongly support the hypothesis [21, 23] that positive feedback is a general mechanism through which epigenetic domains, ultrasensitivity and epigenetic switches might be established and regulated in the cell nucleus. In our case, the conformation-epigenetics feedback plays a major role in the nature and stability of the emerging epigenetic states, which had not previously been appreciated.

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] A. V. Probst, E. Dunleavy, and G. Almouzni, Nat. Rev. Mol. Cell. Biol. **10**, 192 (2009).
[3] B. Strahl and C. Allis, Nature **403**, 41 (2000).
[4] T. Jenuwein and C. D. Allis, Science **293**, 1074 (2001).
[5] B. M. Turner, Cell **111**, 285 (2002).
[6] M. Nicodemi and A. Prisco, Phys. Rev. Lett. **98**, 108104 (2007).
[7] P. Avner and E. Heard, Nature Rev. Genet. **2** (2001).
[8] H. Marks, J. Chow, and S. Denissov, Genome Res. **3**, 1361 (2009).
[9] S. F. Pinter, R. I. Sadreyev, E. Yildirim, Y. Jeon, T. K. Ohsumi, M. Borowsky, and J. T. Lee, Genome Res. **22**, 1864 (2012).
[10] S. Wood and A. Loudon, J. Endocrinol. **222** (2014).
[11] F. Bratzel and F. Turck, Genome Biology **16**, 192 (2015).
[12] A. Angel, J. Song, C. Dean, and M. Howard, Nature **476**, 105 (2011).
[13] L. Hou, D. Wang, D. Chen, Y. Liu, Y. Zhang, H. Cheng, C. Xu, N. Sun, J. McDermott, W. B. Mair, and J.-D. J. Han, Cell Metabolism **23**, 529 (2016).
[14] C. J. Kenyon, Nature **464**, 504 (2010).
[15] H. N. Lim and A. van Oudenaarden, Nature genetics **39**, 269 (2007).
[16] M. Barbieri, M. Chotalia, J. Fraser, L.-M. Lavitas, J. Dostie, A. Pombo, and M. Nicodemi, Proc. Natl. Acad. Sci. USA **109**, 16173 (2012).
[17] C. A. Brackley, S. Taylor, A. Papantonis, P. R. Cook, and D. Marenduzzo, Proc. Natl. Acad. Sci. USA **110**, E3605 (2013).
[18] D. Jost, P. Carrivain, G. Cavalli, and C. Vaillant, Nucleic Acids Research **42**, 1 (2014).
[19] B. Zhang and P. G. Wolynes, Proc. Natl. Acad. Sci. USA **112**, 201506257 (2015).
[20] R. Cortini, M. Barbi, B. R. Care, C. Lavelle, A. Lesne, J. Mozziconacci, and J.-M. Victor, Rev. Mod. Phys. **88**, 1 (2016).
[21] I. B. Dodd, M. a. Micheelsen, K. Sneppen, and G. Thon, Cell **129**, 813 (2007).
[22] K. Sneppen, M. A. Micheelsen, and I. B. Dodd, Mol. Sys. Biol. **4**, 182 (2008).
[23] M. A. Micheelsen, N. Mitarai, K. Sneppen, and I. B. Dodd, Phys. Biol. **7**, 026010 (2010).
[24] I. B. Dodd and K. Sneppen, J. Mol. Biol. **414**, 624 (2011).
[25] N. A. Hathaway, O. Bell, C. Hodges, E. L. Miller, D. S. Neel, and G. R. Crabtree, Cell **149**, 1447 (2012).
[26] K. Sneppen and N. Mitarai, Phys. Rev. Lett. **109**, 100602 (2012).
[27] L. C. M. Anink-Groenen, T. R. Maarleveld, P. J. Verschure, and F. J. Bruggeman, Epigenetics chromatin **7**, 30 (2014).
[28] D. Jost, Phys. Rev. E **89**, 1 (2014).
[29] C. A. Brackley, M. E. Cates, and D. Marenduzzo, Physical Review Letters **111**, 1 (2013).
[30] A. M. Näär, B. D. Lemon, and R. Tjian, Annu. Rev. Biochem. **70**, 475 (2001).
[31] F. Erdel, K. Müller-Ott, and K. Rippe, Ann NY Acad. Sci. **1305**, 29 (2013).
[32] M. C. Barnhart, P. H. J. L. Kuich, M. E. Stellfox, J. A. Ward, E. A. Bassett, B. E. Black, and D. R. Foltz, J. Cell. Biol. **194**, 229 (2011).
[33] G. E. Zentner and S. Henikoff, Nat. Struct. Mol. Biol. **20**, 259 (2013).
[34] H. H. Ng, F. Robert, R. A. Young, and K. Struhl, Mol. Cell **11**, 709 (2003).
[35] T. Garel, H. Orland, and E. Orlandini, EPJ B **268**, 261 (1999).
[36] A. H. F. M. Peters, D. O'Carroll, H. Scherthan, K. Mechtler, S. Sauer, C. Schöfer, K. Weipoltshammer, M. Pagani, M. Lachner, A. Kohlmaier, S. Opravil, M. Doyle, M. Sibilia, and T. Jenuwein, Cell **107**, 323 (2001).
[37] G. Li, R. Margueron, M. Ku, P. Chambon, B. E. Bernstein, and D. Reinberg, Genes Dev. **24**, 368 (2010).
[38] S. Aranda, G. Mas, and L. Di Croce, Sci. Adv. **1**, e1500737 (2015).
[39] E. Lieberman-Aiden, N. L. van Berkum, L. Williams, M. Imakaev, T. Ragoczy, A. Telling, I. Amit, B. R. Lajoie, P. J. Sabo, M. O. Dorschner, R. Sandstrom, B. Bernstein, M. A. Bender, M. Groudine, A. Gnirke, J. Stamatoyannopoulos, L. A. Mirny, E. S. Lander, and J. Dekker, Science **326**, 289 (2009).
[40] D. R. Brumley, M. Polin, T. J. Pedley, R. E. Goldstein, and R. E. Goldstein, J. R. Soc. Interface **12**, 20141358 (2015).
[41] K. Kremer and G. S. Grest, J. Chem. Phys. **92**, 5057 (1990).
[42] A. Colliva, R. Pellegrini, A. Testori, and M. Caselle, Phys. Rev. E **91**, 052703 (2015).
[43] F. Bouchet, S. Gupta, and D. Mukamel, Physica A **389**,

bioRxiv preprint doi: https://doi.org/10.1101/058933; this version posted June 14, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

4389 (2010), arXiv:1001.1479.
[44] T. Garel and H. Orland, EPL (Europhysics Letters) **6**, 307 (1988).
[45] P.-G. de Gennes, J. Phys. (France) Lett. **46**, 639 (1985).
[46] Y. A. Kuznetsov, E. G. Timoshenko, and K. A. Dawson, J. Chem. Phys. **103**, 4807 (1995).
[47] A. Byrne, P. Kiernan, D. Green, and K. a. Dawson, J. Chem. Phys. **102**, 573 (1995).
[48] L. I. Klushin, J. Chem. Phys. **108**, 7917 (1998).
[49] N. Kikuchi, A. Gent, and J. M. Yeomans, EPJ E **9**, 63 (2002).
[50] N. Kikuchi, J. F. Ryder, C. M. Pooley, and J. M. Yeomans, Phys. Rev. E **71**, 1 (2005).
[51] S. Růžička, D. Quigley, and M. P. Allen, Phys. Chem. Chem. Phys. **14**, 6044 (2012).
[52] C. Leitold and C. Dellago, J. Chem. Phys. **141** (2014).
[53] L. A. Mirny, Chromosome Res. **19**, 37 (2011).
[54] S. Redner, Journal of Physics A: Mathematical and General **13**, 3525 (1980).
[55] B. Duplantier, Phys. Rev. B **35**, 5290 (1987).
[56] J. Dekker, K. Rippe, M. Dekker, and N. Kleckner, Science **295**, 1306 (2002).
[57] C. A. Brackley, J. Johnson, S. Kelly, P. R. Cook, and D. Marenduzzo, Nucleic Acids Res. (2016).
[58] K. Müller-Ott, F. Erdel, A. Matveeva, J.-P. Mallm, A. Rademacher, M. Hahn, C. Bauer, Q. Zhang, S. Kaltofen, G. Schotta, T. Höfer, and K. Rippe, Mol. Sys. Biol. **10**, 746 (2014).
[59] D. Michieletto, D. Marenduzzo, and A. H. Wani, arXiv:1604.03041 , 1 (2016).
[60] P. Cook, *Principles of Nuclear Structure and Function* (Wiley, 2001).
[61] T. Garel, H. Orland, and E. Pitard, in *Spin Glasses and Random Fields*, edited by A. Young (World Scientific, 1997) p. 57.
[62] G. Le Treut, F. Képès, and H. Orland, Biophys. J. **110**, 51 (2016).
[63] C. F. Abrams, N. Lee, and S. P. Obukhov, Europhys. Lett. **59**, 391 (2002).
[64] P. M. Chaikin and T. C. Lubensky, *Principles of Condensed Matter Physics* (Cambridge University Press, 2007).

[Supplementary Information]

Simulations Details

The polymer is simulated as a semi-flexible [41] bead-spring chain and in which the beads possess and internal degree of freedom denoted by $q = \{1, 2, 3\}$.

The attraction/repulsion between the beads is regulated by the modified WCA potential as described in the main text:

$$U^{ab}_{LJ}(x) = \frac{4\epsilon_{ab}}{\mathcal{N}}\left[\left(\frac{\sigma}{x}\right)^{12} - \left(\frac{\sigma}{x}\right)^{6} - \left(\frac{\sigma}{x_c^{q_a q_b}}\right)^{12} + \left(\frac{\sigma}{x_c^{q_a q_b}}\right)^{6}\right] \quad \text{for } x \leq x_c^{q_a q_b} \tag{10}$$

and $U^{ab}_{LJ}(x) = 0$ for $x > x_c^{q_a q_b}$. The q-dependent interaction cut-off $x_c^{q_a q_b}$ is set to: (i) $2^{1/6}\sigma$, modeling interaction by steric repulsion only, between beads with different colors, or with color corresponding to no epigenetic marks (i.e., $q = 3$); (ii) $R_1 = 1.8\sigma$ between beads with the same color, and corresponding to a given epigenetic mark (e.g., $q = 1$, or $q = 2$), modeling self-attraction, e.g., mediated by a bridging protein [1]. The free parameter ϵ_{ab} is set so that $\epsilon_{ab} = \epsilon$ for $q_a = q_b = 2$ and $\epsilon_{ab} = k_B T_L$ otherwise. Because the potential is shifted to equal zero at the cut-off, we normalise $U^{ab}_{LJ}(x)$ by $\mathcal{N}$ in order to set the minimum of the attractive part to $-\epsilon$ (see also Fig. S1).

The connectivity is taken into account via a harmonic potential between consecutive beads

$$U^{ab}_{harm}(x) = \frac{k_h}{2}(x - x_0)^2(\delta_{b,a+1} + \delta_{b,a-1}) \tag{11}$$

where $x_0 = 2^{1/6}\sigma$ and $k_h = 200\epsilon$. The stiffness is modelled via a Kratky-Porod term [41]

$$U^{ab}_{KP}(x) = \frac{k_B T_L l_K}{2\sigma}\left[1 - \frac{\boldsymbol{t}_a \cdot \boldsymbol{t}_b}{|\boldsymbol{t}_a||\boldsymbol{t}_b|}\right](\delta_{b,a+1} + \delta_{b,a-1}) \tag{12}$$

where $\boldsymbol{t}_a$ and $\boldsymbol{t}_b$ are the vectors joining monomers a,$a+1$ and b,$b+1$ respectively. The parameters $l_K/2$ is identified with the persistence length l_P of the chain, here set to $l_P = 3\sigma$.

The total potential $U^a(x)$ to which any one bead is then subjected to is the sum over all the possible interacting pairs and triplets, *i.e.*

$$U^a(x) = \sum_{b \neq a}\left(U^{ab}_{LJ}(x) + U^{ab}_{harm}(x) + U^{ab}_{KP}(x)\right). \tag{13}$$

The dynamics of each monomer is evolved by means of a Brownian Dynamics (BD) scheme, *i.e.* with implicit solvent. The Langevin equation reads

$$m\frac{d^2\boldsymbol{r}_a}{dt^2} = -\gamma\frac{d\boldsymbol{r}_a}{dt} - \boldsymbol{\nabla}U^a(x) + \boldsymbol{\xi}_a \tag{14}$$

where γ is the friction coefficient and $\boldsymbol{\xi}_a$ a stochastic noise which obeys the fluctuation dissipation relationship $\langle\xi_{a,\alpha}(t)\xi_{b,\beta}(t')\rangle = 2\gamma k_B T_L \delta_{a,b}\delta(t - t')\delta_{\alpha,\beta}$, where the Latin indexes run over particles while Greek indexes over Cartesian components. Using the Einstein relation $D = k_B T_L/\gamma = k_B T_L/3\pi\eta\sigma$ where η the solution viscosity one can define a "Brownian" timestep $\tau_{Br} = \sigma^2/D$ as the time required for a bead to diffuse its own size. The dynamics is then evolved using a velocity-Verlet integration within the LAMMPS engine in Brownian dynamics mode (NVT ensemble).

The systems are simulated in a box of linear size L and in dilute regime (assuming each monomer occupying a cylindrical volume $\pi\sigma^3/4$ one can estimate the volume fraction as $\rho = M\pi\sigma^3/4L^3 \simeq 0.01\%$). The box is surrounded by a purely repulsive wall in order to avoid self-interactions through the periodic boundaries. The initial configuration is typically that of an ideal random walk in which the beads assume a random type q. We then run 10^4 τ_{Br} timesteps in which only a increasingly stronger steric soft repulsion is considered between every pair of beads, while their type is left unaltered. The exact form of the soft potential we use is

$$U^{ij}_{soft}(d) = A\left[1 + \cos\frac{\pi d_{ij}}{d_c}\right] \tag{15}$$

where $d_c = 2^{1/6}\sigma$ is the cutoff distance and A the maximum of the potential at $d_{ij} = 0$. This pre-run "warming-up"

bioRxiv preprint doi: https://doi.org/10.1101/058933; this version posted June 14, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

drives the ideal random walk conformations to ones having self-avoiding statistics as it removes the overlaps between monomers without generating numerical "blow-ups".

From the configurations thereby generated we start the production run of 10^6 τ_{Br} timesteps in which M writing moves are performed every 10^3 τ_{Br} timesteps in order to allow the chains to explore enough surrounding space before performing a new recolouring move. Each recolouring move is accepted or rejected using a Metropolis algorithm, *i.e.* the acceptance probability is given by

$$p(q \to q') = \min\left(1, e^{-\Delta E/k_B T_{\rm Rec}}\right), \tag{16}$$

where ΔE is the difference between the new and the old energies and, more importantly, the temperature that appears in the exponent is the "recolouring" temperature $T_{\rm Rec}$, not necessarily identical to T_L used in the Langevin equation for the stochastic noise.

One can map the simulation units into real units by considering $\sigma = 30$ nm $\simeq 3$ kbp, which amounts to around 20 nucleosomes made by 150 bp. The total polymer length is therefore $M = 2000\sigma \simeq 6\ 10^4$ nm or 6 Mbp. In addition, by considering the solution viscosity that of water, $\eta \simeq 10$ cP and room temperature $T = 300$ K finally leads to a Brownian time $\tau_{Br} = 3\pi\eta\sigma^3/k_B T_L = 0.6$ ms. The recolouring move is performed every $10^3\tau_{Br} = 0.6$ seconds, which is within the range of the biological timescales for the action of histone-modifying enzymes. The total simulation runtime is set to 10^3 recolouring steps, which is equivalent to 10 minutes in real time (XXX).

The Detailed Balance is broken when $T_P \neq T_L$.

According to the Kolmogorov criterion, in a dynamics satisfying detailed balance the product of the transition rates over any closed loop of states must be the same as its own reverse. This is not the case when $T_{\rm Rec} \neq T_{Langevin}$. In fact, one can imagine a simple case when two beads with the same q and contacting each other are kicked apart by a thermal fluctuation. This happens with probability $p^q_{\rm near\to far} = \exp\left(-\epsilon/k_B T_L\right)$. At this stage, a change in the bead type (q) happens with probability 1, as there is no energy penalty. When the beads have different type they can bump into each other with probability again 1, as there is no attraction or penalty in being near one another (as long their distance is greater than $2^{1/6}\sigma$). Now attempting the swap that brings the two beads to have the same q also happens with probability 1 as this move is energetically favourable. Therefore we obtain

$$p_{loop} = \exp -\frac{\epsilon}{k_B T_L}. \tag{17}$$

By performing the loop the other way round (*i.e.* change q first, then get far apart, change back q, and finally fall back into the well) one instead obtains

$$p_{loop^{-1}} = \exp -\frac{\epsilon}{k_B T_{\rm Rec}} \neq p_{loop}. \tag{18}$$

The two are equal only if $T_{\rm Rec} = T_L$. In particular, if $T_L > T_{\rm Rec}$ the "direct" loop is more likely to happen than its reverse, while the opposite is true if $T_L < T_{\rm Rec}$.

Second Virial Coefficient

It is straightforward to extract the second virial coefficient u_2 using the Mayer relation and eq. (10) [62]:

$$u_2^{ab} = -\int d^3x \left(e^{-\beta U^{ab}_{LJ}(x)} - 1\right). \tag{19}$$

We find that u_2^{ab} is positive (u_2^{rep}) for $q_a \neq q_b$ and is instead negative (u_2^{att}) when $q_a = q_b$. In particular, we find that $u_2^{rep} \simeq 4.396$ while u_2^{att} ranges from -9.3 (for $\epsilon = 1k_B T_L$) to -400 (for $\epsilon = 5k_B T_L$).

Contact Maps – 2 State model

In this section (Fig. S2) we report a series of contact maps for the "two-state" model, from the moment of the quench. One can notice that while for high interaction parameters, the folding of the polymer, as well as the network of interactions, is frozen, for values of α close to the transition point $alpha_c = 1$, the contact map instead evolves into a full checker-board interaction pattern.

First -Order-Like Nature of the Transition

In Fig. S3 we report a series of pictures representing the joint probability distribution $P(R_g, \tilde{m})$, *i.e.* the probability of observing the system in a certain state with given signed magnetisation $\tilde{m}$ and radius of gyration R_g. As discussed in the text, the system undergoes a transition from a swollen (large R_g), disordered ($\tilde{m} \sim 0$) to a compact (small R_g) and ordered (coherent magnetisation $\tilde{m} \simeq \pm 1$). Here we show

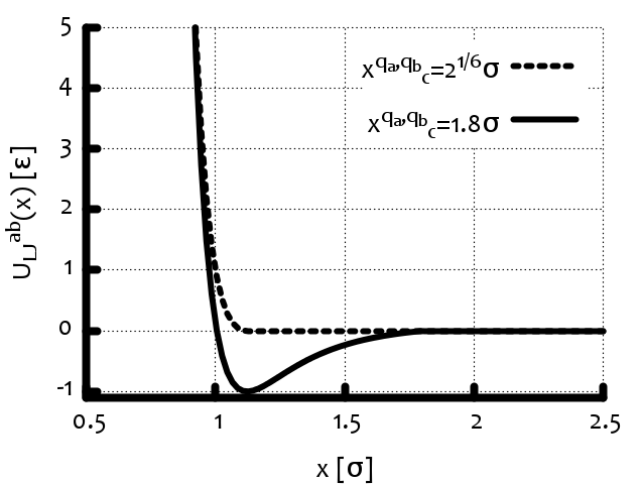

Figure S1. Shape of the modified WCA potential for cut-off $x_c^{q_a,q_b} = 1.8\sigma$ (when $q_a = q_b$) and $x_c^{q_a,q_b} = 2^{1/6}\sigma$ (when $q_a \neq q_b$)

bioRxiv preprint doi: https://doi.org/10.1101/058933; this version posted June 14, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

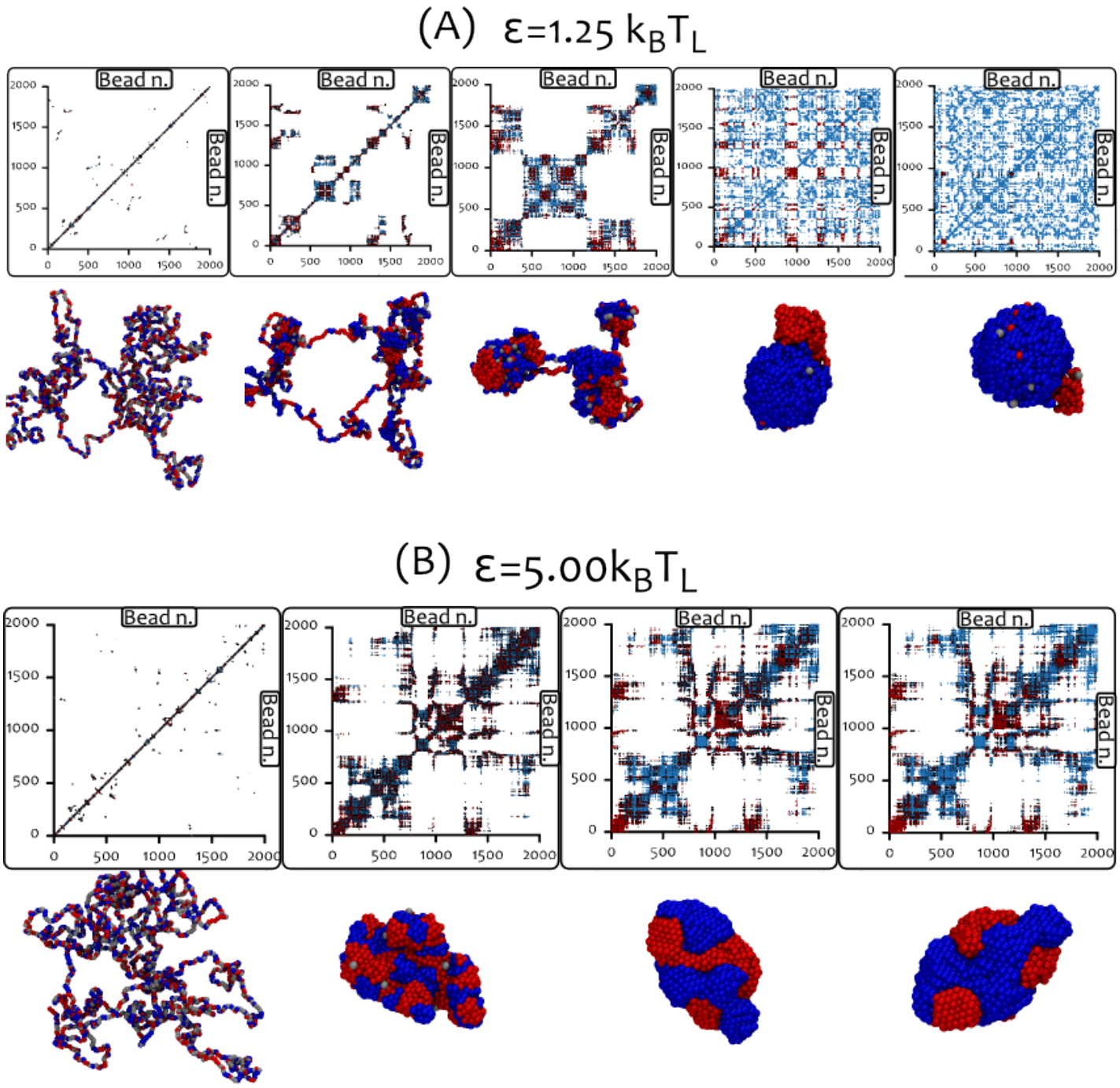

Figure S2. **Contact maps for the "two-state" model**. In this figure we report the contact maps for the snapshots shown in the main text Fig. **2**. As one can notice, while low $\alpha = \epsilon/k_B T_L$ leads the system to a checker-board contact map at large times, high values of α, or deep quenches, freeze the network of contacts. The colors in the contact maps are determined as follow: red (blue) (grey) if the contact is between two red (blue) (grey) beads, and black otherwise, *i.e.*, for mixed contacts. One can notice that high values of α force the chain to display a large number of mixed contacts; this is due to the fact that there is a large gain in free energy even with few coherent contacts. These mixed contacts are then observed to slowly re-arrange locally, *i.e.* without displaying large changes in the contact map, but only swapping black dots with colored ones. This is in turn consistent with the measurement of the exchange dynamics reported in Fig. S10: this is in fact slow and glass-like for high α.

that the transition is first-order-like, in that there is a value of temperature (interaction energy ϵ) for which the system shows the coexistence of both phases, *i.e.* the free energy has three minima. In order to show this, we needed to sample the phase space near the critical point $\alpha_c = \epsilon/k_B T_L|_c$ as broadly as possible. We therefore performed 100 independent simulations of systems with a polymer $M = 50$ beads long and runtime 10^6 τ_{Br} recording the state (epigenetic magnetisa-

bioRxiv preprint doi: https://doi.org/10.1101/058933; this version posted June 14, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

tion and radius of gyration) every 10^3 τ_{Br} for several values of $\epsilon \in [k_BT_L, 1.5k_BT_L]$. As one can notice from Fig. S3, the system undergoes a discontinuous goes from a swollen and disordered phase for $\epsilon \leq 1.1k_BT_L$ to an ordered and compact one for $\epsilon \geq 1.2k_BT_L$. We observe that considering longer chains lowers the critical point towards $\epsilon \simeq k_BT_L$, but we did not probe the value of α_c as accurately as in this case. We have instead investigated hysteresis cycles with a chain with $M = 2000$ beads by taking a collapsed configuration generated evolving the chain under $\alpha = 1$, and reducing the interaction parameter to $\alpha = 0.9$. In Suppl. Move M7 we compare the behaviour of two chains starting from the two (ordered and disordered) phases and subject to the same $\epsilon = 0.9k_BT_L$ for $10^6\tau_{Br}$ timesteps, therefore supporting the coexistence of two stable phases in this range (see also snapshots in Fig. S4).

Finally, we highlight that we do not observe switching between the two symmetric stable states, *i.e.* $\tilde{m} = +1$ and $\tilde{m} = -1$, for a chain with $M = 2000$ beads, but only for shorter chains (see Fig. S5 and Suppl. Movie M8). Such a switching was reported in the literature for effectively 1D models [21, 26, 28], which also considered small number of nucleosomes.

This result is compatible with the fact that switching occurs when the system overcomes the energy barrier between the two states, which grows with both, the interaction strength ϵ and the number of intrachain interactions (which increases with M). In other words, the average first passage time from one state to the other can be predicted by a Kramers formula,

$$\langle \tau \rangle \sim \frac{k_BT_L}{Dw_ow_d} e^{M\epsilon/k_BT_L}, \tag{20}$$

where, D is the diffusion coefficient of the system in the magnetisation space, and $w_{o,d}$ the oscillation frequencies in the potential at the ordered (o) and disordered (d) coordinates in the magnetisation space. Given that the energy barrier between the minima is proportional to the size of the polymer, it follows that the probability of switching decays exponentially with M.

Decay of the Radius of Gyration

In this section we illustrate a simple physical reasoning to rationalise the exponential decay of the gyration radius during the collapse at the transition point. Although there are some authors who argue that the collapse should be self-similar in time, and therefore, following a power law [45, 63], we have not found evidence of this self-similar collapse. This fact is presumably be due to either the finite size of the chain used in our investigation, or to the initial condition, as in our simulations we start from configurations far from a stretched coil, which is instead the situation often considered in theoretical models [45]. Therefore in our case the common assumption of neglecting long-ranged loops at the early stages of the collapse [45] may not be appropriate. Apart from the theory explored in Ref. [46], we have not found in the literature a simple argument as to why the size of the polymer should decrease exponentially in time during the collapse. For this reason we illustrate a simple argument below.

If one takes the growth (in number of monomers) of the pearls at very early times as $g \sim t^\beta$, with β unknown for the moment, the volume of the pearls will grow as

$$R_p^d \sim g^{d\nu} \sim t^{\beta\nu d} \tag{21}$$

since each pearl is a crumpled globule $\nu = 1/d$ and hence

$$R_p^d \sim g \sim t^\beta \tag{22}$$

the total number of monomers in pearls is gN_p (where N_p is the number of pearls), therefore the number of inter-pearl monomers (not in the pearls) is

$$N_{ip} = N - gN_p \sim N\left(1 - g\frac{N_p}{N}\right) \tag{23}$$

$$\sim N\left(1 - \frac{N_p}{N}t^\beta\right) \sim Ne^{-N_pt^\beta/N} \tag{24}$$

as at early times $gN_p/N \ll 1$ and t is small by definition of "early-time". When pearls begin to appear, they are separated by a 3D distance given by the average number of inter-pearl monomers to the exponent ν and in particular the 3D distance R_{ip} is

$$R_{ip} \sim \left(\frac{N_{ip}}{N_p}\right)^\nu \sim \left(\frac{N}{N_p}\right)^\nu e^{-\nu N_pt^\beta/N}. \tag{25}$$

For $t = 0$, Eq. (25) correctly predicts that the typical size of inter-pearl distance is the whole polymer (as $N_p = 1$). For $t \neq 0$, it predicts a stretched exponential decay of the gyration radius for $\beta < 1$, and a simple exponential, for $\beta = 1$. Therefore our argument provides a reason for a non-power-law decay of R_g.

We note that this argument is valid at very early times, or when the chain is large enough that the number of monomers belonging to the growing pearls N_p is much smaller than the number of monomers in the chain. It does not make any assumption regarding the presence of long range loops, while it makes the assumption that segments of the polymer not sucked in by the pearls are still in a statistically relaxed conformation ($R_{ib} \sim N_{ib}^\nu$). Although we have observed that the growing of pearls introduce competing tensions along the chain, at early times (or for very large chains), such forces do not spread across the whole chain, therefore leaving intra-blobs segments, tension-free.

Although we cannot give an estimation for β within our reasoning, this is not needed to prove the exponential decay of R_g in time during the collapse. This exponent might assume values in between $\beta = 1$ for a mean-field dynamics of a conserved order parameter [64] to $\beta \simeq 0.66$ as observed numerically for the coarsening of pearls during a homopolymer collapse [47]. A more detailed study of the early stages of the collapse dynamics of a recolourable polymer (or annealed copolymer) might shed some light into the precise value of β for this case, and on the precise nature of the decay of the radius of gyration (stretched versus simple exponential).

bioRxiv preprint doi: https://doi.org/10.1101/058933; this version posted June 14, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

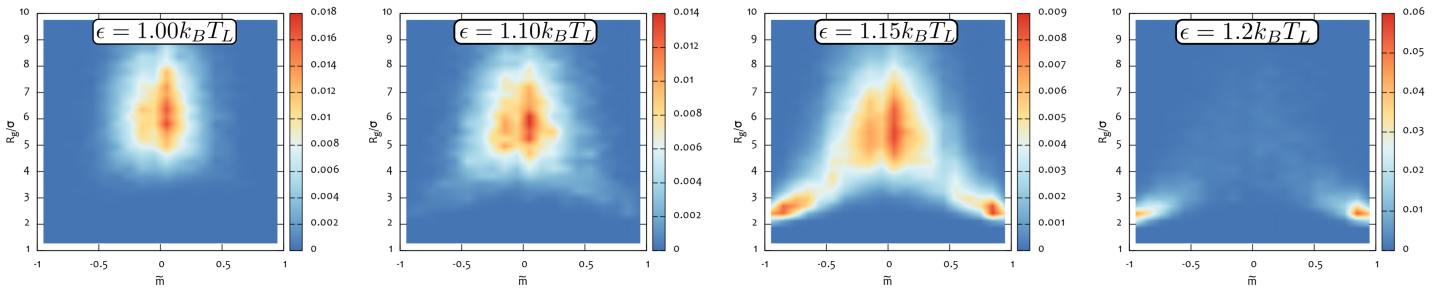

Figure S3. **First-order-like transition for the two-state model for a polymer with** $M = 50$**.** In this figure we report the probability $P(R_g, m)$ for a chain $M = 50$ beads long to display a given radius of gyration R_g and signed epigenetic magnetisation $\tilde{m}$ for four values of the interaction parameter $\alpha = \epsilon/k_B T_L$ near the critical point. Each one of the pictures is done by averaging over the dynamics of 100 independent simulations for a chain $M = 50$ beads long and left run for 2 10^6 τ_{Br}. We note that longer chains display lower values of the critical point $\alpha_c \simeq 0.95 - 1$, although we did not thoroughly explored the phase space for longer chains.

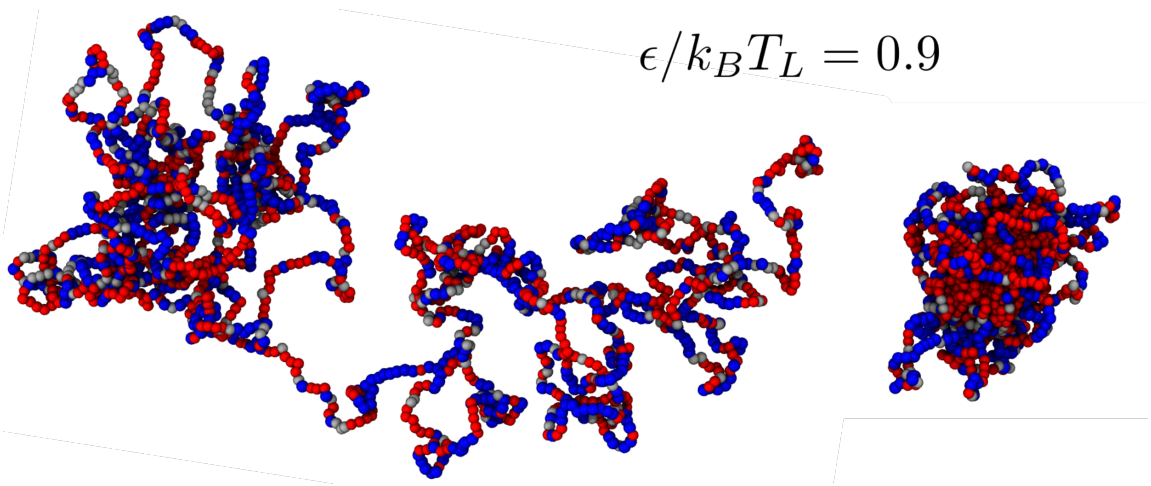

Figure S4. **Coexistence and Hysteresis.** Once that a chain with $M = 2000$ beads has gone through the transition by using $\epsilon/k_B T_L = 1$, we reduce the interaction parameter to $\epsilon/k_B T_L = 0.9$ and compare the behaviour of two chains starting from the two phases subject to this same temperature for long time (10^6 τ_{Br}). We find (as the figure shows) that there is coexistence between a disordered swollen phase (top) and the compact ordered one (bottom). Furthermore, the way we probed the coexistence also suggests the presence of hysteresis cycles characteristic of discontinuous phase transitions (see main text).

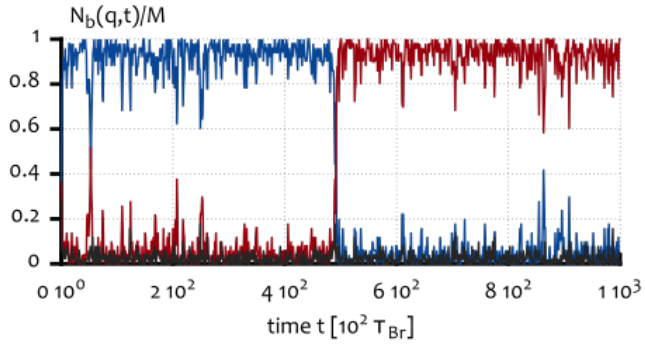

Figure S5. **Switch-like behaviour is observed in short polymers**. In agreement with previous findings we here recover a switch-like behaviour in the case of short polymers. Here we simulated a polymer $M = 50$ beads long with interaction strength $\epsilon = 1.25 k_B T_L$. (see Suppl. Movie M8)

Contact Probability

In this section we report (in Figs. S6 and S7) the contact probabilities measured from our simulations. In order to highlight the differences with mean field models we measured the contact probability of a single bead (index $b = 1000$ along the chain, in Fig. S6). While one would observe a contact probability $P_c(m) \sim m^{-c}$ for an effective 1D Ising model with long range interactions tuned by c, for any chosen bead in the system, here we observe that the contact probability assumes a shape closer to a sum of delta-functions. This suggests a strong preferential selection of certain contacts along the polymer and a strong deviation from a mean-field type of interaction pattern. In Fig. S6 we report our findings by plotting (in the left hand side graphs), a symbol for every bead that is in contact with bead $b = 1000$ at each time-step and for $\alpha = 1.25 - 5$ $k_B T_L$. These plots show that the "fuzziness" that characterises mobile networks of contacts disappears when $\alpha \geq 3 k_B T_L$. At these values of interaction strength, the contacts between some beads are

bioRxiv preprint doi: https://doi.org/10.1101/058933; this version posted June 14, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

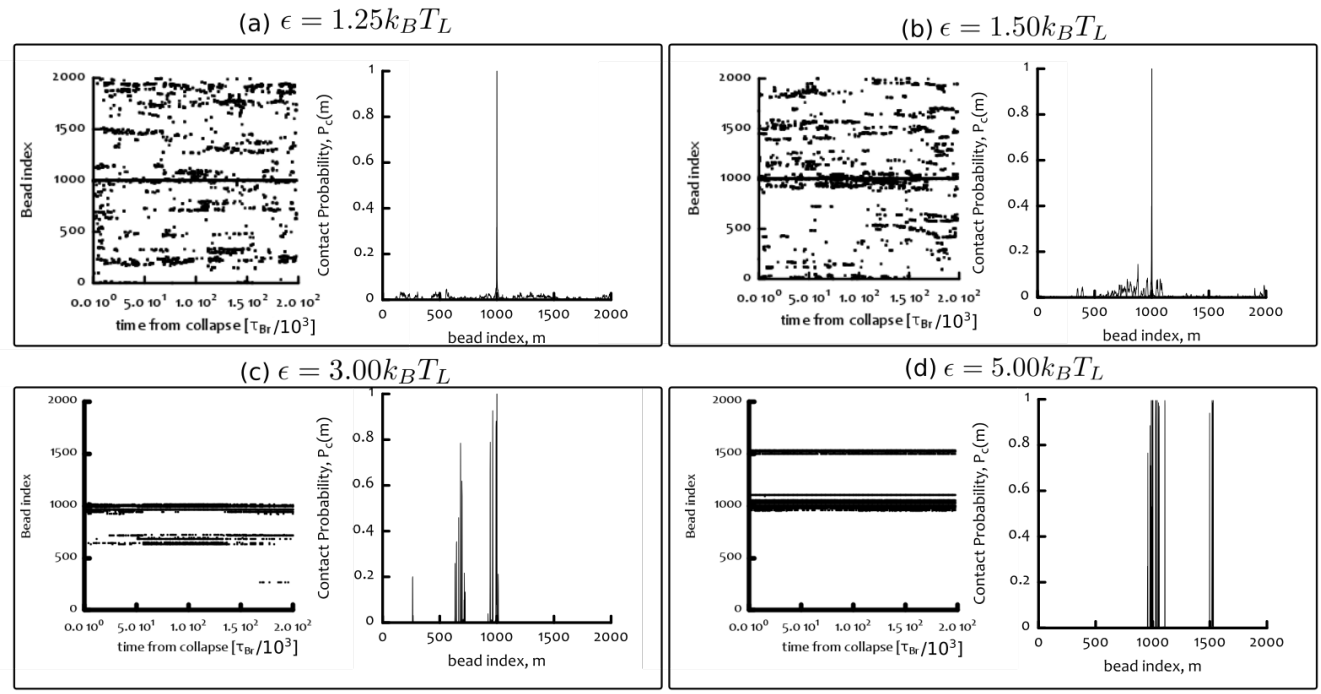

Figure S6. **The folding of the polymer creates a quenched network of contacts.** In this figure we show four pairs of graphs, each one for a different choice of α. For each pair, the left plot is made by placing a point every time a bead i is in contact with bead $b = 1000$ at a certain time-step after te collapse. As one can notice, these plots are highly dynamic (or "fuzzy") at low α, this is because the network of contact is rearranging quickly during the simulation. Higher interaction strengths instead induce the selection of a subset of all possible interactions and create a frozen network of contacts. The right plots in fact show the averaged (over time) contact probability for the bead $b = 1000$. While for low α the plots show a decay that might be well captured by a power law, for interaction strengths $\alpha > 3$ they display a strong departure from a power-law decay of the contact probability $P_c(m)$ which is instead strongly peaked around a subsection of all contacts.

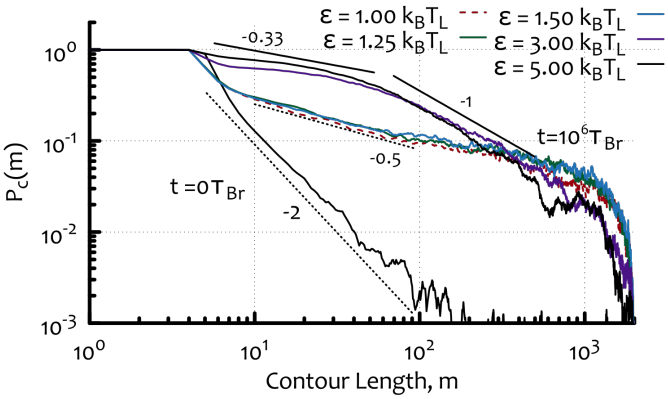

Figure S7. **The contact probability averaged over independent replicas and beads leads to a power law statistics of contacts.** Here we show that a more standard contact probability emerges when it is averaged over different simulation replicas and monomers. The initial configurations show a very steep decay compatible with the self-avoiding walk statistics $c \simeq 2$ (for an ideal random walk one would have $c = d/2 = d/2$) while the collapsed states show $c \simeq 0.5$ for $1.0 \geq \epsilon/k_BT_L = \alpha \leq 1.5$. Higher interaction parameters lead to an enhancement of local contacts ($c = 1/3$) followed by a steeper decay $c = 1$ at longer ranges compatible with the fractal globule conjecture.

present at all times and they never exchange.

The plots in Fig. S6 also show (on the right hand side) the time averaged contact probability (again for bead $b = 1000$). The graphs capture the strong departure from a mean-field-like interactions for high interaction strengths as in fact $P_c(m)$ resemble a sum of delta-functions rather than a power-law function. The picture that emerges is therefore similar to that of spin-like variables interacting on a network, where the edges are established by the collapse dynamics. When the interaction parameters are higher than a certain value the edges of the network are frozen in place, resembling a spin glass. We finally stress, that although we observe this departure from the mean-field assumption, the average of $P_c(m)$ over many beads and many replicas of the system gives a more "traditional" power-law decay as we show in Fig. S7. In particular, we find $P_c(m) \sim m^{-c}$ with c ranging from 1/3 to 1/2 for different interaction strengths at the end of the collapse dynamics (see Fig. S7), while they all start from a situation where $c \gtrsim 2$ compatible with a self-avoiding walk statistics (an ideal random walk would have $c = d/2 = 1.5$).

Connectivity

In this section we report several quantities to characterise the change in network connectivity. As described in the main text, we track the average number of neighbours $N_n(t)$, and

bioRxiv preprint doi: https://doi.org/10.1101/058933; this version posted June 14, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

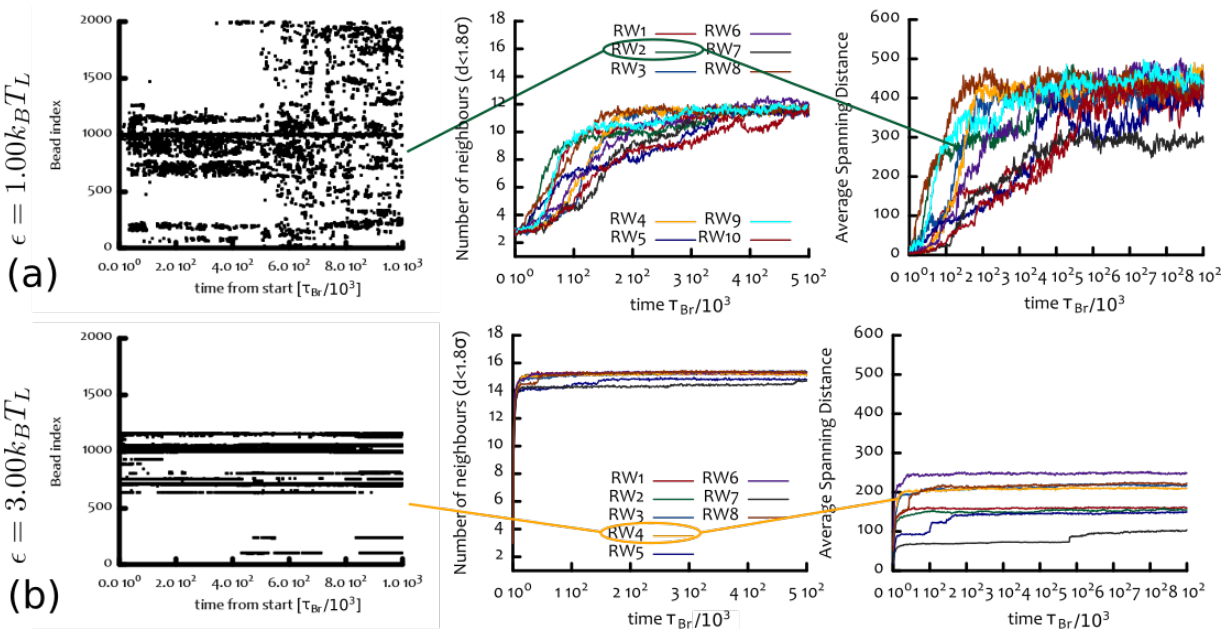

Figure S8. **Long-range interactions appear dynamically during the collapse. (a)** and **(b)** show (from left ot right) the time-dependence of the interaction plot, number of neighbours and average spanning distance for three choices of α and various simulations (denoted by "RW" and the index of the simulation). The choices of α correspond to **(a)**$\epsilon = k_B T_L$ and **(b)** $\epsilon = 3k_B T_L$. This figure shows that long range contacts develop dynamically during the collapse for α near the transition value α_c, while for higher values of α the interactions are frozen (*i.e.* the do not evolve in time) and although the number of neighbours is larger, the spanning distance is shorter, ultimately leading to slower epigenetic dynamics (see main text).

also the average spanning distance $\Delta_s(t)$ defined as

$$\Delta_s(t) = \frac{1}{M} \frac{\sum_{a \neq b} |a-b| P_{ab}(t)}{\sum_{a \neq b} P_{ab}(t)}. \tag{26}$$

The dynamical changes of these quantities during the collapse of various systems (denoted with "RW" and the index of the simulation) are reported in Fig. S8. From this figure it is important to notice that while there is an evident increase in number of neighbours and spanning distance for $\alpha = \epsilon/k_B T_L = 1$, the same is not observed for higher interaction parameters. In these cases, *e.g.*, the case with $\alpha = 3$, the network of interactions is frozen, the number of neighbours quickly saturates to the maximum value and the spanning distance is frozen and achieves a steady-state a smaller values than that achieved for lower values of α.

The average of these quantities across replicas are then shown in Fig. S9(a)-(b). Once again, one can readily see that while the number of neighbours increases and then plateaus at large α, the spanning distance has a more complex dependence on the interaction strength. In Fig. S9(c) we also show the time averaged values (taken after the collapse and over the last 5 10^5 τ_{Br}) alongside the value of the radius of gyration as a function of α. One can notice that both Δ_s and R_g are non-monotonic functions of α, therefore suggesting the existence of a critical α_c at which the response of the system changes by further increasing of interaction strength. In particular one can notice that for $\alpha \geq 3$ the spanning distance starts to decrease and the radius of gyration to increase, corresponding to the formation of more short ranged network and more frustrated configurations at higher α.

Finally in Fig. S9(d) we report the value of the neighbour exchange rate κ_n divided by the average number of neighbours N_n at any time after the collapse: it can be seen that κ_n/N_n always reaches a steady state. This steady state value monotonically decreases and, in particular, undergoes a sharp transition around $\alpha_c \simeq 3$ (see Fig. S10), above which the network of interactions is quasi-frozen.

Contact Maps – Two-State Model with Intermediate State

In this section we report, in Fig. S11, the contact maps for the "two-state" model with intermediate state (IS) for different time-steps and during the artificial local recolouring. We started from the mixed metastable state (MMS) and artificially recolored a small (10%) segment in the middle of the chain. From Fig. S11 one can notice that this localised perturbation quickly drives the system toward the epigenetically coherent and globular state. As we show in the main text, this ordered phase is instead robust against major global reorganisation events such as a semi-conservative replication. This ultra-sensitive response, *i.e.* a dramatically different response of the system to an external stimulus depending on the current state of the system, can be appreciated by looking at the large re-organisation and phase transition driven by such a small perturbation (Fig. S11 and main text Fig. **5**) as opposite to the preserved coherent state even after a replication event (Fig. **5** at late time).

bioRxiv preprint doi: https://doi.org/10.1101/058933; this version posted June 14, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

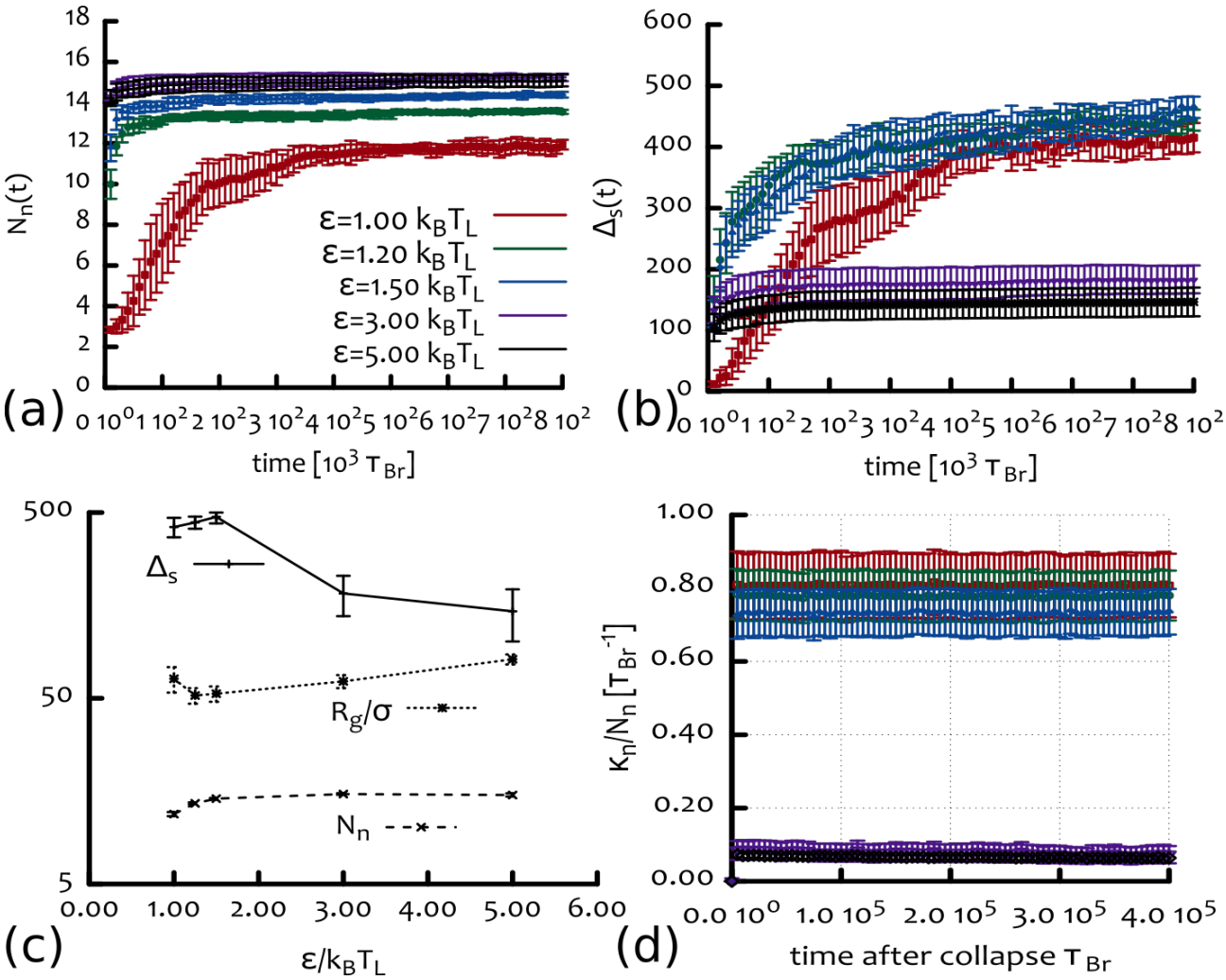

Figure S9. **Short ranged and glassier networks of contacts emerge at large interaction strengths.** **(a)** Number of neighbours per bead, $N_n(t)$, averaged over different simulations. **(b)** Average spanning distance $\Delta_s(t)$ averaged over independent replicas. **(c)** The values (averaged over last 5 10^5 τ_{Br}) of the number of neighbours, $\langle N_n \rangle$, and spanning distance $\langle \Delta_s \rangle$ plotted together with the average radius of gyration $R_g(t)/\sigma$ as a function of α. **(d)** Fraction of exchanged neighbours per Brownian time κ_n/N_n is here shown to achieve a steady state after the collapse.

Contact Maps – Two-State Model with Broken Detailed Balance

In this section we report the contact maps for the "two-state" model with broken detailed balance at different timesteps within the averaged window. In the main text we show the average contact map generated by summing over instantaneous contact maps that look as the ones reported in Fig. S12. As one can notice, the out-of-diagonal contacts are temporary and very mobile. Averaging over such frames lead to the "block-like" structure reported in the main text and strongly reminiscent of the TAD-like structures often reported by capture experiments in eukaryotic cells.

Theta point for an homopolymer

In the main text we describe an argument for which our "two-state" model with broken detailed balance leads to TADs, *i.e.* a block-like pattern along the contact map. One of the key point of the argument is that the theta temperature for a homopolymer is larger than that of a heteropolymer (or copolymer). This means that one can start from the disordered swollen state, in which the polymer resembles a random copolymer rather than an homopolymer, and set an interaction strength $\epsilon = k_B T_L$ such that a random copolymer would not collapse, but for which an equivalently-long homopolymer would instead become a globule. In this way, when larger 1D epigenetically coherent domains appears along the polymer, due to the lowering of the "recolouring" temperature T_{Rec}, temporary globules shorter than the whole chain, can be marginally stable.

In order to approximately estimate the value of theta tem-

bioRxiv preprint doi: https://doi.org/10.1101/058933; this version posted June 14, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

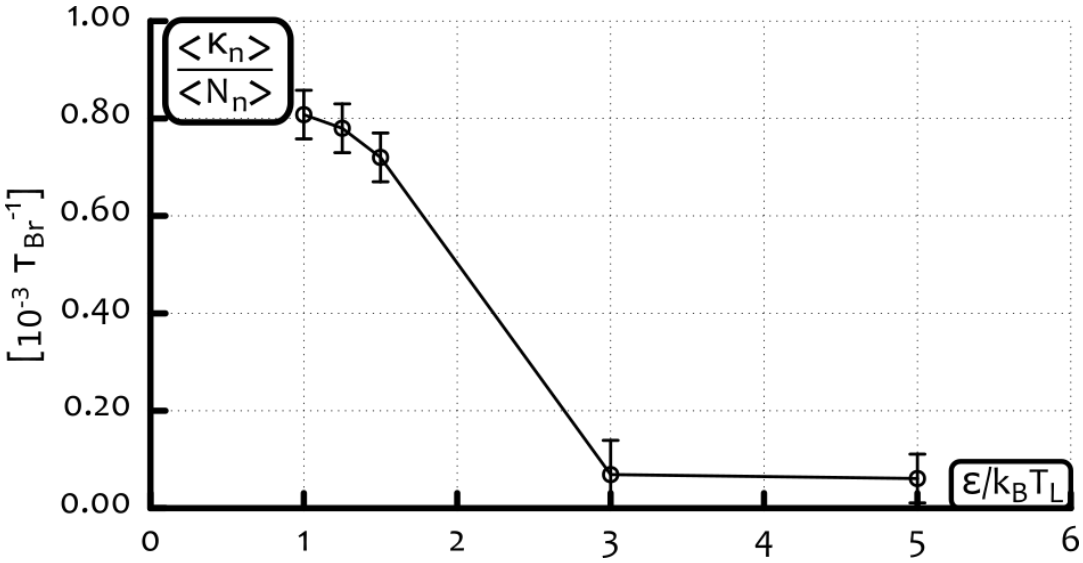

Figure S10. **The exchange of neighbours dramatically slows down at high α's.** In this figure we report the behaviour of κ_n/N_n as a function of $\alpha = \epsilon/k_BT_L$. Thus, we show the fraction of neighbours exchanged per time-step on average by each bead. One can readily note a dramatic change in the rate κ_n when $\alpha \geq 3k_BT_L$, for which case it attains a value of almost zero. This implies that the network of interactions is frozen, corresponding to a glassy exchange dynamics.

perature for a homopolymer simulated within our framework we report the behaviour of the radius of gyration for a homopolymer $M = 2000$ beads long for different temperatures T_L and starting from a swollen self-avoiding configuration (Fig. S13). The curves are averaged over 10 independent realisations of the system. We observe that for $T_L \gtrsim 1.85\epsilon/k_B = T_L^{*,h}$ the polymer is not smaller than its initial state, and therefore consider this temperature as roughly the theta-point for the homopolymer (non-recolouring) case. At this stage it is worth reminding that the critical point for an equally long "recolorable" polymer is at $T_L^{*,r} \simeq \epsilon/k_B$.

The finding reported in this section are in support to the fact that in the two-state model with broken detailed balance, we have used $T_L = 1.75\epsilon/k_B$, therefore much larger than $T_L^{*,r}$ but lower than $T_L^{*,h}$. For this reason, epigenetic domains that are large enough can temporarily aggregate into globules.

The Epigenetic Correlation Length

In analogy to the Ising 1D magnetic correlation length, the "epigenetic correlation length", can be measured by imagining that each bead has a epigenetic state q that can take any of the possible states: red, blue or grey. The epigenetic correlation function can therefore be expressed in terms of these variables as

$$g_{1D}(m) \equiv \langle \delta_{q(0),q(m)} \rangle = \left\langle \frac{1}{M-m} \sum_{ij} \delta_{q_i q_j} \delta(|i-j|-m) \right\rangle, \tag{27}$$

and where the average is performed over independent system replicas and over different (uncorrelated) times. Since analogous to the 1D Ising (or Potts) magnetisation correlation function, $g_{1D}(m)$ takes the functional form

$$g_{1D}(m) \sim e^{-m/\xi(T_{\rm Rec})} \tag{28}$$

where the correlation length $\xi(T_{\rm Rec})$ diverges with M as the system breaks the symmetry between red/blue epigenetic states. When the system displays different coexisting epigenetic states (multi-state regime), either in the glassy phase of the simplest "two-state" model or in the stable "block-like" organised regime of the model with broken detailed balance, $\xi(T_{\rm Rec})$ is finite.

This length-scale is clearly dependent on $T_{\rm Rec}$ as in fact, even in the case the polymer was an immobile straight line, the correlation length would increase and ultimately diverge as $T_{\rm Rec} \to 0$. Furthermore, in our model the dynamics of the polymer is coupled to the epigenetic organisation, and therefore $\xi(T_{\rm Rec})$ is expected to depend more subtly on $T_{\rm Rec}$. For this reason we measured $g_{1D}(m)$ for several cases and reported its decay in Fig. S14. In particular, we compare $g_{1D}(m)$ for different values of $T_{\rm Rec}$ (with fixed $T_L = 1.75\epsilon/k_B$) and for the case that leads to glass-like dynamics with $T_{\rm Rec} = T_L = 0.2\epsilon/k_B$ (*i.e.* $\alpha = 5$). While the latter has a very low absolute value of $T_{\rm Rec}$, the decay of $g_{1D}(m)$ is much (more than one order of magnitude) sharper and this is due to the very fragmented domain structure that we have shown for the glass-like case in the main text and can be readily be observed in the kymograph of Fig. **2**. On the other hand, for fixed $T_L = 1.75\epsilon/k_B$ and varying $T_{\rm Rec}$ one clearly recovers the increase in $\xi(T_{\rm Rec})$ which is around 200 monomers for the specific case $T_{\rm Rec} = 0.1\epsilon/k_B$ (see Fig. S14). This value also determines the size of the domains observed in Fig. **6** (as 2000 beads are divided into ~ 9 domains, 5 red and 4 blue, giving an average domain size of roughly 220 beads).

These findings support the argument described in the last section of the main text, where we argued that the com-

bioRxiv preprint doi: https://doi.org/10.1101/058933; this version posted June 14, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

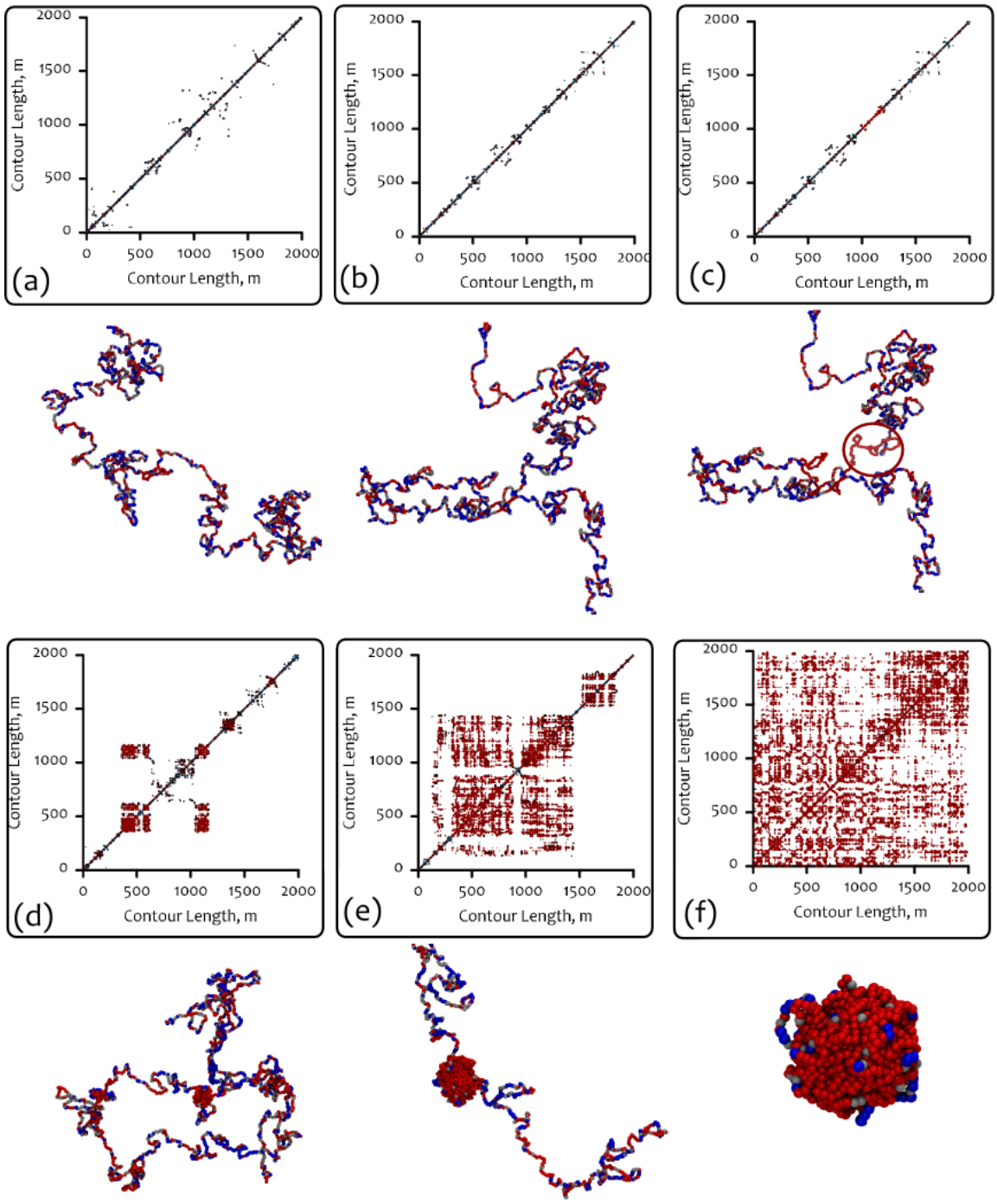

Figure S11. **Contact maps and snapshots for the "two-states" with intermediate state model before and after the perturbation.** In this figure we report the evolution of the contact map for the two-state with intermediate state model during the artificial recolouring. **(a)-(c)** refer to **(a)** 4 10^5 τ_{Br} before the perturbation while **(b)** and **(c)** show the contact map immediately before and after the recolouring (in **(c)** one can in fact notice a small red segment in the middle of the contact map). **(d)-(f)** refer to **(d)** 2 10^5 τ_{Br}, **(e)** 4 10^5 τ_{Br} and **(f)** 6 10^5 τ_{Br} after the perturbation. From this figure one can readily appreciate the dramatic change in conformation (phase transition) driven by the localised artificial recolouring.

petition between extensive 1D epigenetic coherent domains and 3D disordered (swollen) polymer configurations could generate temporarily stable epigenetically coherent globules without taking over the whole chain. This is therefore done by tuning the system at $T_L^{*,r} < T_L < T_L^{*,h}$ (see previous section) and, importantly, by tuning $T_{\rm Rec}$ so as to generate 1D extensive coherent domains in agreement with the behaviour of the correlation length of a 1D Ising-like system.

bioRxiv preprint doi: https://doi.org/10.1101/058933; this version posted June 14, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

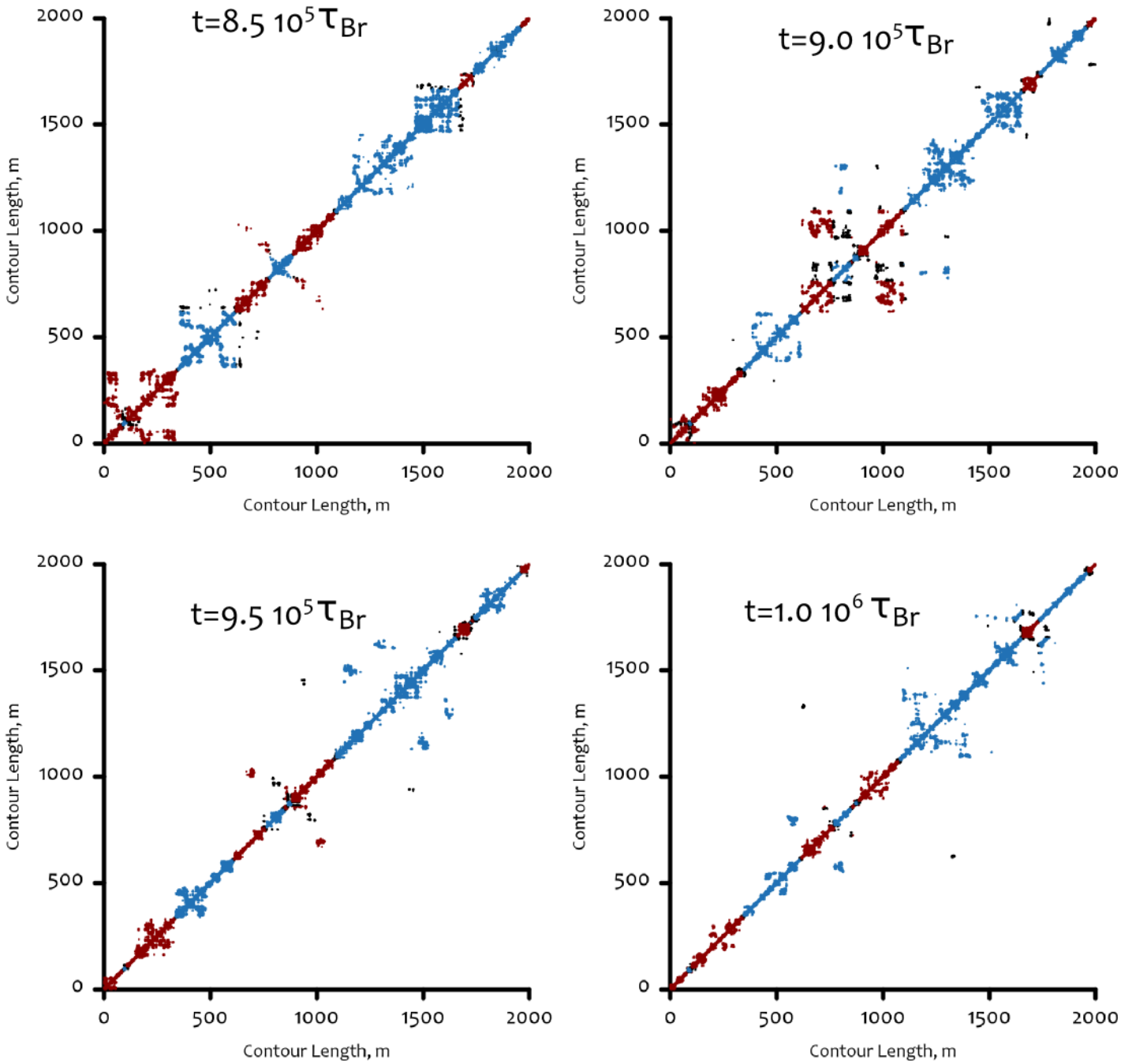

Figure S12. **Snapshots of the contact maps for the "2 states model" with broken detailed balance model**. This figure shows the instantaneous contact maps at different timesteps within the averaged window. As in the case of capture experiments, the averaged contact map reveals more clearly the organisation of the TADs which are less visible from the single snapshots. Nonetheless, from this figure one can appreciate the temporary aggregation of globules, whose underlying 1D epigenetic segment is preserved during the simulation.

Other regimes of the "two-state" model with broken detailed balance

In this section we report other regimes that we observe in the "two-state" model with broken detailed balance. In particular, this time we fix $T_{\text{Rec}} = 0.1\epsilon/k_B$ and change the value of T_L in the range $T_L \in [1.25, 2]\epsilon/k_B$. We observe that $T_L \leq 1.5\epsilon/k_B$ leads to a collapse of the coil into a single-state dominated globule, similar to that observed in the standard "two-state" model, but with a higher critical temperature T_L. On the other hand, $T_L \geq 1.75\epsilon/k_B$ leads to a more interesting regime, where local globules coexist along the chain, giving the averaged contact map a characteristic "block-like". We also observe that higher temperatures T_L

bioRxiv preprint doi: https://doi.org/10.1101/058933; this version posted June 14, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

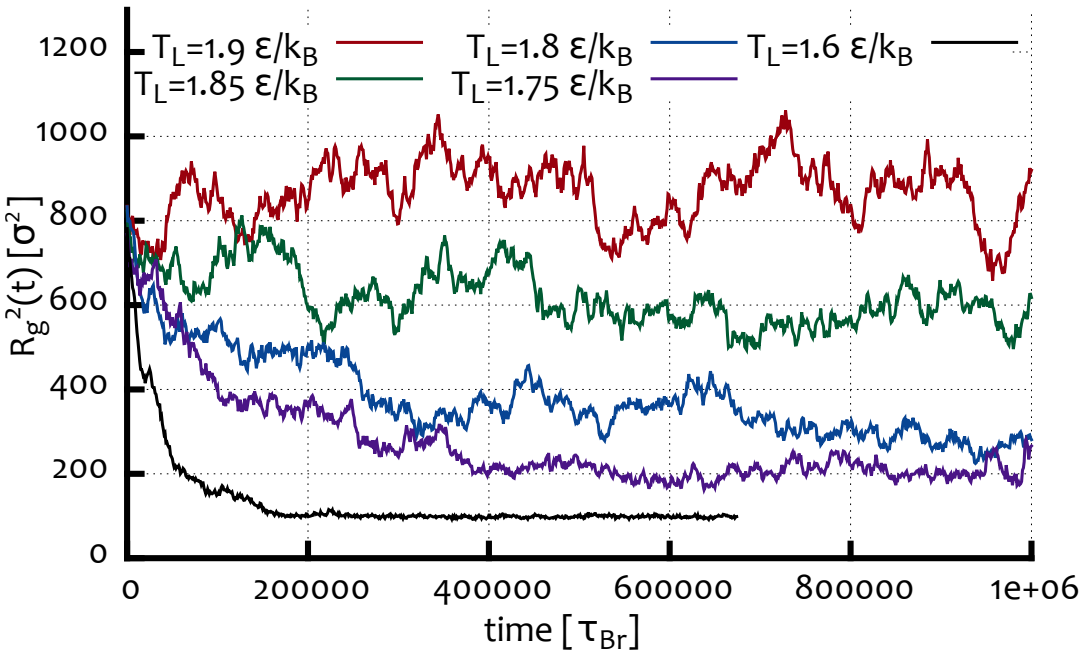

Figure S13. **Theta-point for a non-recolouring homopolymer** $M = 2000$ **beads long**. This figure shows the behaviour of the radius of gyration (squared) for a non-recolouring homopolymer starting from a self-avoiding swollen configuration. The curves are averaged over 10 independent replicas of the system. Temperatures above $TL \simeq 1.85\epsilon/k_B$ do not seem to drive the collapse of the polymer and we therefore interpret this as a rough indication of the theta-point of the system.

lead to a smaller size of the domains for fixed and small T_{Rec}.

bioRxiv preprint doi: https://doi.org/10.1101/058933; this version posted June 14, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

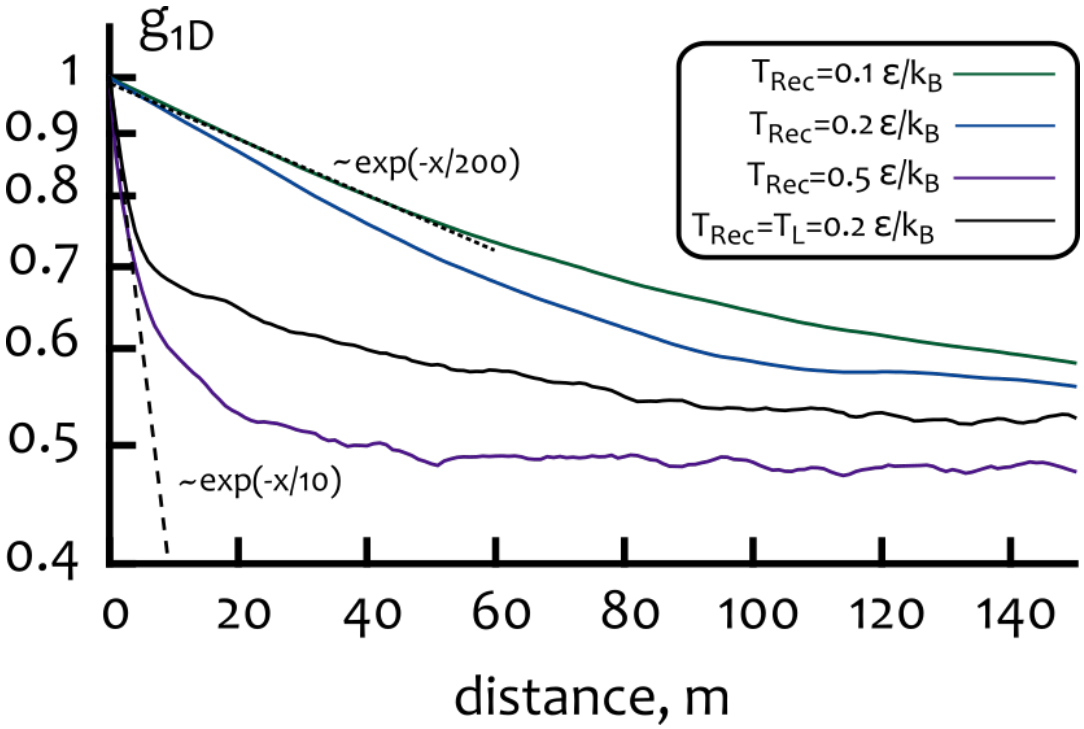

Figure S14. **The epigenetic correlation length**. This figure shows (in a log-linear plot) the correlation function $g_{1D}(m)$ for different T_{Rec} and $T_L = 1.75\epsilon/k_B$ (colored lines) and for $T_{\mathrm{Rec}} = T_L = 0.2\epsilon/k_B$ (black line) which we have shown to lead to a state with multiple domains and glass-like dynamics. As expected the correlation length $\xi(T_{\mathrm{Rec}})$ increases when T_{Rec} is lowered, similarly to the case of 1D Ising system. The long-length behaviour of the correlation length is $g_{1D}(m \to \infty) \sim 1/n_{\mathrm{states}}$ which is 1/2 since only two out of three states are actually self-attracting.

bioRxiv preprint doi: https://doi.org/10.1101/058933; this version posted June 14, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

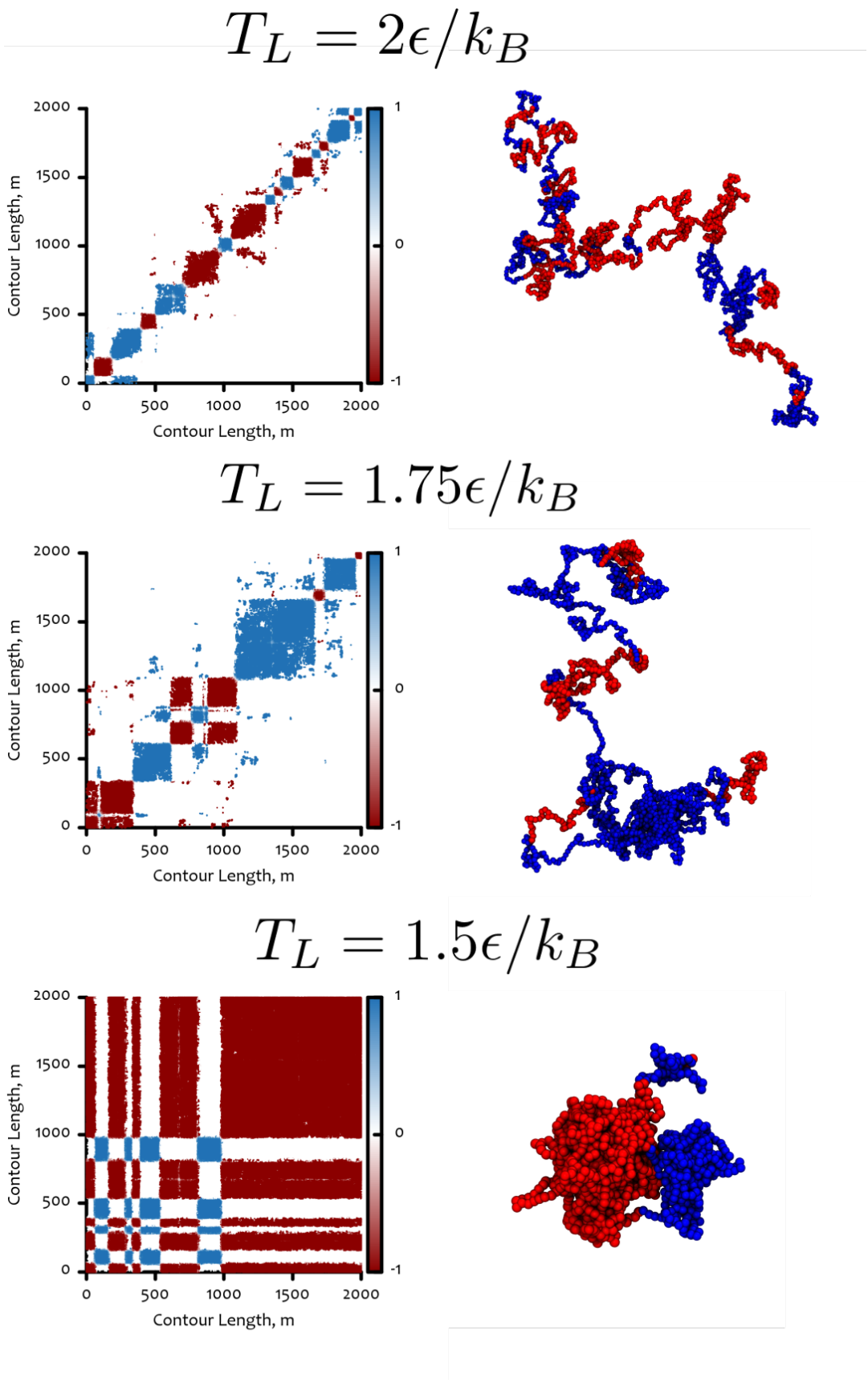

Figure S15. **Classification of the regimes in the "two-state" model with broken detailed balance**. This figure shows the observed regimes for the two-state model with broken detailed balance. Here, we fix $T_{\mathrm{Rec}} = 0.1\epsilon/k_B$ and vary (from bottom to top) $T_L \in \{1.5, 1.75, 2\}\epsilon/k_B$. We observe that $T_L \leq 1.5\epsilon/k_B$ leads to a collapse of the coil into a single-state dominated globule, similar to that observed in the standard "two-state" model, but with a higher critical temperature. On the other hand, $T_L > 1.5\epsilon/k_B$ lead to a more interesting regime, where local globules coexist along the chain giving a characteristic "block-like" pattern to the averaged contact map.