Conformation Regulation of the X Chromosome Inactivation Center: A Model

Antonio Scialdone[1], Ilaria Cataudella[2], Mariano Barbieri[3], Antonella Prisco[4], Mario Nicodemi[5]*

1 Department of Computational and Systems Biology, John Innes Centre, Norwich, United Kingdom, 2 Center for Models of Life, Niels Bohr Institute, Copenhagen, Denmark, 3 Dipartimento di Scienze Fisiche, Università di Napoli "Federico II," INFN, Napoli, Italy, 4 CNR Istituto di Genetica e Biofisica "B. Traverso", Napoli, Italy, 5 Dipartimento di Scienze Fisiche, Università di Napoli "Federico II," INFN, CNR-SPIN, Napoli, Italy

Abstract

X-Chromosome Inactivation (XCI) is the process whereby one, randomly chosen X becomes transcriptionally silenced in female cells. XCI is governed by the *Xic*, a locus on the X encompassing an array of genes which interact with each other and with key molecular factors. The mechanism, though, establishing the fate of the X's, and the corresponding alternative modifications of the *Xic* architecture, is still mysterious. In this study, by use of computer simulations, we explore the scenario where chromatin conformations emerge from its interaction with diffusing molecular factors. Our aim is to understand the physical mechanisms whereby stable, non-random conformations are established on the *Xic's*, how complex architectural changes are reliably regulated, and how they lead to opposite structures on the two alleles. In particular, comparison against current experimental data indicates that a few key cis-regulatory regions orchestrate the organization of the *Xic*, and that two major molecular regulators are involved.

Introduction

X-Chromosome Inactivation (XCI) is the vital process occurring in female mammalian cells whereby one randomly selected X is transcriptionally silenced to balance dosage with respect to males [1–4]. XCI is regulated by a region on the X chromosome, the X inactivation center (*Xic*), which encompasses a key group of neighboring non-coding genes (see Fig. 1.A) including, e.g., *Jpx*, *Xist*, *Tsix* and *Xite* [1–4]. The fate of the X is determined by its *Xist* gene which is strongly upregulated on the future inactive X and repressed on the other X. In turn, *Xist* is negatively regulated by *Xite/Tsix*, and positively regulated by *Jpx*, *Rnf*12, and other factors [5–7].

Before random XCI starts, a complex epigenetic program, coupling transcription and chromatin remodelling [8,9] to pluripotency factors [10–12], produces a state where the *Xic* has the same spatial conformation on the two X chromosomes [13] and both *Xist* alleles are just weakly active. Upon XCI, an unknown symmetry breaking mechanism determines the opposite behaviour of the two *Xist*, and induces alternative modifications of the three-dimensional conformation of their *Xic* [13,14]. Finally, on the designated inactive X further chromatin reorganizations occur as a heterochromatic compartment forms into which genes are recruited to be silenced [3,15]. Several molecular factors are known to be involved in the process [3,4], including noncoding transcripts, chromatin modifiers and organizers, such as CTCF (a Zn finger having arrays of binding sites on the *Xic*), Dnmt3a, Oct4 and other pluripotency factors [9–12,16,17]. Different models have been proposed to describe random XCI [18–22], but still none to elucidate its associated chromatin changes, whose nature remains mysterious.

To understand the principles of chromatin organization, within the murine *Xic* case study, here we explore the scenario where chromatin conformations emerge from its interaction with diffusing molecular factors. We discuss general physical mechanisms whereby random Brownian molecules can: *i)* succeed in establishing stable, non random conformations on the chromosomes; *ii)* reliably regulate specific conformational changes; and *iii)* produce opposite transformations on identical alleles exposed to the same environment ("symmetry breaking"). We investigate by computer simulations a schematic model consisting of two identical polymers which interact with a concentration of diffusing molecules (see Fig. 1.B). In the light of current *Xic* 3C data [13], the model poses that along each polymer three types of regions exist type-α, β and γ) and predicts the existence of two types of regulatory molecules (type-A and B).

We show that the system thermodynamic stable states fall in distinct classes corresponding to different conformations. The polymers spontaneously select one of them according to molecule concentration/binding energy. Conformational changes are driven by thermodynamic phase transitions which act switch-like, regulated by given concentration/binding energy thresholds. The two polymers are exposed to the same environment, yet they can undergo alternative architectural modifications: we show that a symmetry breaking mechanisms is activated if the homotypic interaction between regulatory molecules rises above a threshold.

Comparison to experimental observations [1–5,13,21] suggests that the regions envisaged by the model can be approximately mapped along the *Xic* sequence as illustrated in Fig. 1.B, while type-A and B complexes could be related to an activating and a blocking regulator of *Xist*.

Author Summary

In mammal female cells X-Chromosome Inactivation (XCI) is the vital process whereby one X, randomly chosen, is silenced to compensate dosage of X products with respect to males. XCI is governed by a region on the X, the X Inactivation Centre (*Xic*), which undergoes a sequence of conformational modifications during the process. The two *Xic* are exposed, though, to the same environment, and it is obscure how they attain different architectures. By use of computer simulations of a molecular model, here we individuate general physical mechanisms whereby random Brownian molecules can assemble chromatin stable architectures, reliably regulate conformational changes, and establish opposite transformations on identical alleles. In the case-study of the murine *Xic*, our analysis highlights the existence of a few key regulatory regions and molecular factors. It also predicts, e.g., the effects of genetic modifications in the locus, which are compared with current deletion/insertion experiments. The physical mechanisms we describe are rooted in thermodynamics and could be relevant well beyond XCI.

Model

We represent the relevant region of each X chromosome (see scheme in Fig. 1.B) by a standard model of polymer physics, a self-avoiding bead chain [23]. In the light of *Xic* current 3C data [13], we pose that along each polymer there are, for simplicity, two type-α regions which have an array of binding sites for type-A
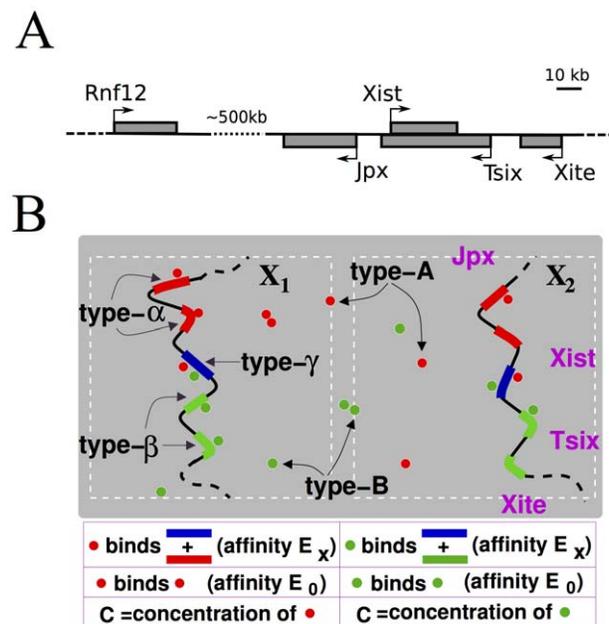
Figure 1. The model. Panel A is an illustration of the region of the X Inactivation Centre (*Xic*) around the *Xist* gene. The scheme in **panel B** zooms on the key regions of the two polymer model investigated here. Each polymer has two type-α (red), two type-β (green) and a type-γ (blue) regions. Type-α and type-γ can be bridged by type-A molecules (red circles) with an affinity E_X; type-β and type-γ by type-B molecules (green circles). Each molecular species has a concentration c. Type-A molecules have also a homotypic mutual interaction of affinity E_0, and similarly type-B ones. The presumptive mapping areas on the *Xic* are also illustrated (right panel).
doi:10.1371/journal.pcbi.1002229.g001

Brownian molecular factors. Each polymer has also two type-β regions with binding sites for a different kind of molecular factors (type-B). Finally, the polymers have a type-γ region whose binding sites can be bound by either type-A or B molecules. Thus, type-A molecules (resp. type-B) can bridge a type-α (resp. type-β) and a type-γ site. For simplicity, with no loss of generality, we consider the case where the two types of molecules have the same concentration, c, and the same affinity, E_X, for all binding regions. Similarly, we assume that type-α and type-β regions have the same number of binding sites, n_0, than type-γ. The value of n_0 is fixed to have a total binding site number of the order of known *Xic* binding molecules. As CTCF is a general chromatin organizer which has been associated to XCI and its *Xic* binding sites have been well characterized [17], we use it as an example (and set $n_0 = 20$). For simplicity, n_0 is here also the length of the intervening inert sequences between them. Type-A (resp. type-B) molecules can bind, with multiple valency, each other with affinity E_{AA} (resp. E_{BB}); we set $E_{AA} = E_{BB} \equiv E_0$ and, considering the number of binding domains of CTCF, the valency to four.

We investigate by Monte Carlo (MC) simulations the conformations of the system as they spontaneously emerge when the three control parameters, (c, E_X, E_0), are varied. For computational purposes, the system lives in a cubic lattice with a lattice spacing d_0, whose value corresponds to the typical size of a DNA binding site, and can be roughly estimated to be $d_0 \sim 10 nm$. The volume concentration of molecules in our model, c, can be related to molar concentrations ρ: $\rho \sim c/d_0^3 \mathcal{N}_A$, $\mathcal{N}_A$ being the Avogadro number (details in Text S1). Thus, for instance, a typical nuclear protein concentration of $\rho \sim 0.1 \mu mole/litre$ would correspond to $c \sim 10^{-2}\%$. Below we consider concentrations in the range $c \sim 10^{-4} - 10^{0}\%$ and binding energies in the weak biochemical scale (a few units in $k_B T$). Finally, conversion of MC time unit to real time is obtained by imposing that the diffusion constant of our polymers is of the order of measured chromatin diffusion constants (see Text S1 for details).

Results

Establishing stable interactions

We first show that diffusing molecules can produce a looped conformation on each polymer where type-α and type-β stably interact with type-γ region. The process is based on a thermodynamic mechanism (a phase transition, in the thermodynamic limit) which acts switch-like when concentration/affinity of binding molecules rise above a threshold [24].

Before describing our MC results in details, we illustrate the underlying mechanisms. A single, say, type A molecule forms a bridge between type-α and type-γ regions via the stochastic double encounter of the molecule with its binding sites. This is, though, an unlikely event, especially if molecule concentration, c (or E_X, see below), is small. And the half-life of such a bridge is short when weak biochemical interactions are considered. Thus, on average the regions float away from each other (see pictorial representation in the bottom panel of Fig. 2, "Open State"). At higher c (or E_X), however, many a molecule can bind type-α/γ regions and stabilize the conformation via a positive feedback mechanism as their bridges reinforce each other and facilitate the formation of additional bridges. The concentration where such a positive feedback mechanism starts winning marks the threshold above which stable contacts are established (pictorial representation in the bottom panel of Fig. 2, "Stable Interaction").

This pictorial scenario summarizes our MC results. For sake of simplicity, we consider first the case where molecule mutual interaction is turned off, $E_0 = 0$, and set as initial configuration of
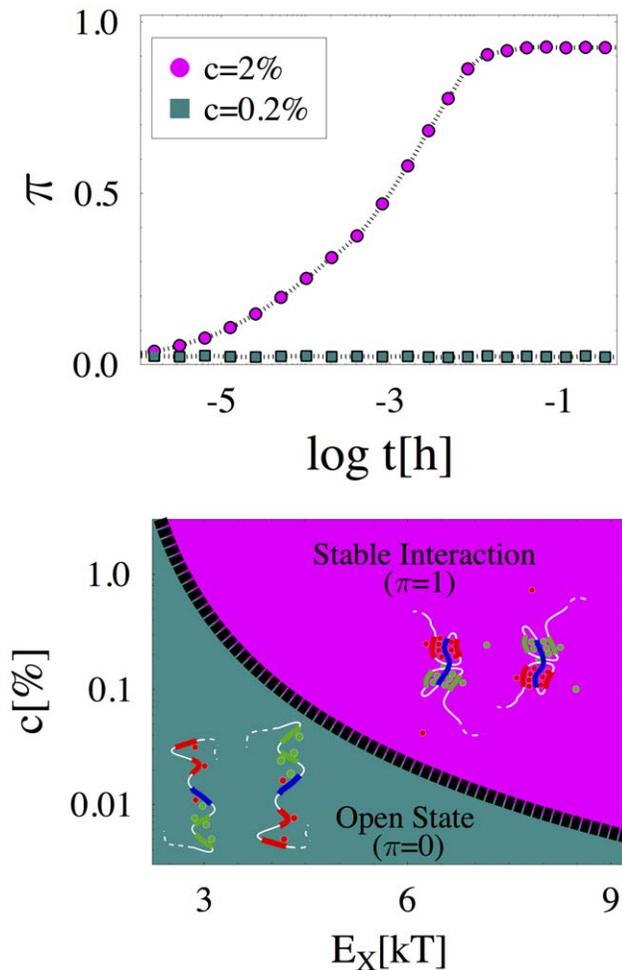
Figure 2. Conformational Switches and the establishment of stable interactions. Top Panel π is the probability of type-α and type-β regions to loop stably onto type-γ region. Its time evolution, from an initial open polymer conformation (see schematic representation in the Bottom Panel, "Open State"), is shown for two characteristic values of the concentration, c (here $E_X = 3kT$ and $E_0 = 0$). For $c = 0.2\%$, π is zero at all times: neither type-α nor type-β regions succeed in forming stable contacts with type-γ, and the polymer conformation remains open. For $c = 2\%$, after a transient of the orders of minutes, π approaches one: a stable, looped conformation is established (see schematic picture in the Bottom Panel, "Stable Interaction"). **Bottom Panel** The conformation phase diagram in the (E_X, c) plane is shown (for $E_0 = 0$): in the region below the sharp transition line, $c_{tr}(E_X)$ (black dashed line), the polymers are found in an open state; above $c_{tr}(E_X)$, they exhibit a conformation change, symmetrical on the two polymers, as a stable interaction of type-α and type-β with type-γ region is established.
doi:10.1371/journal.pcbi.1002229.g002

Conformation switch and sharp regulation

In the space of the control parameters, (c, E_X), a sharp line separates the two regimes, as shown in Fig. 2 bottom panel: when c or E_X are small, contacts cannot be stable and $\pi = 0$; conversely, above the transition line the two loops conformation is reliably established on each polymer, and $\pi = 1$. Such a line marks the boundary between two thermodynamic phases [25]: it corresponds to the point where the entropy loss due to loop formation is compensated by the energy gain obtained from the establishment of the corresponding bridges.

The discovery of such a switch-like behaviour can also explain how loop formation can be sharply and reliably regulated in the cell by increasing the concentration of specific molecular mediators or the affinity to their DNA target sites, e.g., by chromatin or molecule modifications.

The position of the transition line is also dependent on the number of available binding sites, n_0, since, schematically, the overall binding energy scale is $n_0 E_X$. Thus, non-linear threshold effects in genetic deletion/insertions of the locus exist.

Threshold values in real nuclei

From Monte Carlo results we can predict concentration (or energy) thresholds in real nuclei. For instance, *in vitro* measures of CTCF DNA binding energies give $E_X \sim 20kT$, a typical value for TFs [26,27]: an extrapolation from Fig. 2 then predicts a threshold $c_{tr} \sim 10^{-3}\%$, corresponding to a typical nuclear protein molar concentration $\rho \sim 10^{-2} \mu mole/litre$ (see Text S1).
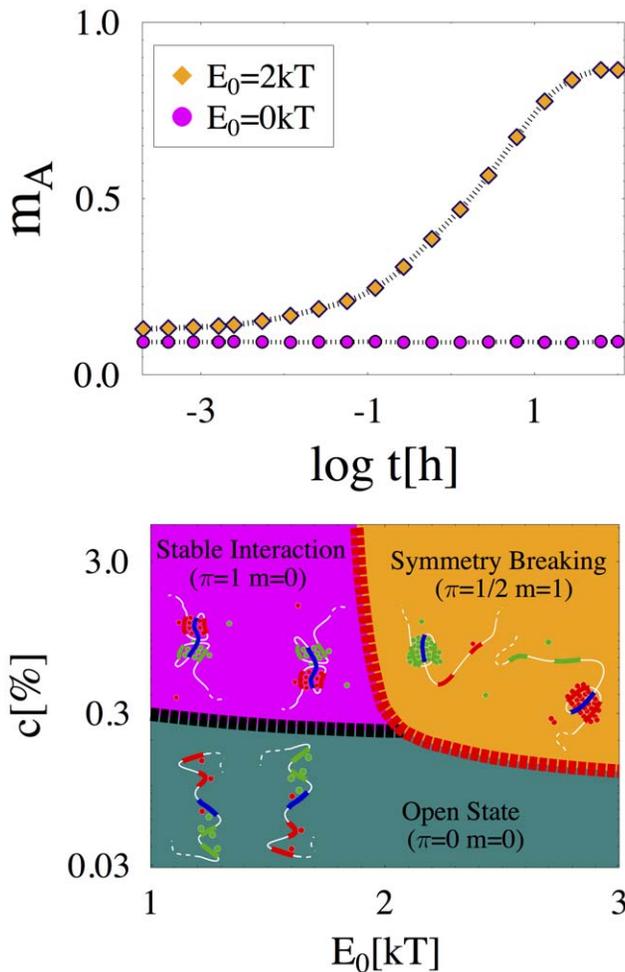
Finally, the mechanism leading to stable loop formation has to be fast enough to serve functional purposes. In our model we find that stable interactions are established on scales of the order of minutes (see Fig. 2 top panel and Text S1), a range consistent with biological expectations.

Symmetry Breaking mechanism

The mechanism to induce conformational changes illustrated above acts "symmetrically" on the two polymers. Now we show that molecule homotypic interaction, E_0, can break the polymer symmetry via a different thermodynamic mechanism. More precisely, if E_0 (and c, see below) is above a critical threshold, a single major aggregate of type A molecules and a single one of type B are formed because of homotypic binding cooperativity: in facts, the energy gain in forming a single cluster of A/B molecules (which maximizes the number of possible chemical bonds) compensates, if E_0 is large enough, the corresponding entropy reduction. The single, say, type A aggregate will then randomly bind just one polymer, leaving the other one "naked" (pictorial representation in the bottom panel of Fig. 3, "Symmetry Breaking").

Type-A and B aggregates bind opposite polymers because A and B molecules compete for binding sites in the type-γ region. Hence, if a fluctuation increases the presence of, say, A molecules on one polymer, cooperativity tends to favor their assembling at that site and B molecules are expelled; in turn, the depletion of A around the other polymer favors the assembling of B molecules on it. On the polymer where the A cluster binds the type-γ region, the B-related loci can no longer be stably linked, and their loop opens; the opposite situation happens on the other polymer.

The above scenario results from our MC simulations. We measured the symmetry breaking order parameter, $m_A = |\rho_A^{(1)} - \rho_A^{(2)}| / (\rho_A^{(1)} + \rho_A^{(2)})$, where $\rho_A^{(i)}$ is the average local concentration of A molecules around the type-γ region of polymer $i = 1, 2$. The m_A parameter is close to zero if an equal amount of A molecules is present around the two polymers, whereas it approaches one if the symmetry is spontaneously broken (m_B and

the polymers a randomly open conformation. We measure the interaction order parameter, $\pi = (p_A + p_B)/2$, where p_A (resp. p_B) is the probability to have, on a polymer, a contact of a type-α (resp. type-β) with type-γ region. If neither type-α nor type-β regions are in contact with γ, the order parameter is zero, $\pi = 0$; if only one pair is stably interacting then $\pi \simeq 1/2$; finally, $\pi = 1$ if both type-α and type-β loops are established. Fig. 2 top panel shows the MC time evolution of $\pi(t)$ for two values of c: if c is small, π remains indefinitely close to zero, $\pi \sim 0$, as no stable contact is statistically possible; instead, if c is high enough, π grows to a value close to one, $\pi \simeq 1$, showing that both the type-α and β loops are formed.

Figure 3. The Symmetry Breaking (SB) mechanism. Top Panel The SB parameter, $m_A = |\rho_A^{(1)} - \rho_A^{(2)}|/(\rho_A^{(1)} + \rho_A^{(2)})$, is the normalized average difference of type-A molecule density around type-γ region of polymers 1 and 2. Its dynamics, from the initial symmetrical polymer looped state (as in the schematic picture in the Bottom Panel, "Stable Interaction"), is shown for two characteristic values of molecule homotypic interaction energy, E_0 (here $c=2\%$ and $E_X=3kT$). If $E_0=0$, m_A is close to zero: molecules are equally distributed around the polymers. If $E_0=2kT$, after a transient of about ten hours, m_A approaches one, i.e., either $\rho_A^{(1)} \to 0$ or $\rho_A^{(2)} \to 0$: molecules have aggregated around only one of the polymers, and their binding symmetry is broken (as in the schematic picture in the Bottom Panel, "Symmetry Breaking"). **Bottom Panel** The phase diagram in the (E_0,c) plane (for $E_X=3kT$) has three phases. If E_0 is below the transition line, $E_{tr}^{SB}(c)$ (red dashed line), the system is in one of its symmetric phases: the "Open State" phase (at low c) or the symmetrical "Stable Interaction" phase. If $E_0 > E_{tr}^{SB}(c)$, the conformational symmetry is broken ("Symmetry Breaking" phase): the type-α loop persists only on one randomly chosen polymer, and type-β on the other.
doi:10.1371/journal.pcbi.1002229.g003

$m \equiv (m_A + m_B)/2$ behave analogously). Fig. 3 top panel shows the time evolution of $m_A(t)$ from an initial configuration corresponding to the symmetric state (schematic picture in the bottom panel of Fig. 3, "Stable Interaction") where each polymer has two stable loops as seen before: if E_0 is small, m_A remains close to zero at all times and the system remains in a symmetric state; conversely, if E_0 is high enough, m_A approaches one because A molecules reside mostly around just one, randomly chosen polymer and the symmetry is broken (schematic picture in the bottom panel of

Fig. 3, "Symmetry Breaking"). The phase diagram of Fig. 3 bottom panel shows that the symmetry breaking mechanism is switch-like too: in the (c,E_0) space, as soon as a narrow transition line is crossed the system switches from a symmetrical polymer state to a broken polymer symmetry state. More details are in the Text S1.

For sake of simplicity, we considered the case where the concentration/DNA affinity of molecules A and B are the same. However, such an assumption does not affect our general results. The only condition for the Symmetry Breaking and Configurational Switch mechanisms to be triggered is that concentration/interaction energy of both types of molecules rise above the appropriate threshold.

Symmetry Breaking in real nuclei

As far as XCI is concerned, the predicted single B molecule aggregate is interpreted as an *Xist* repressing factor (a Blocking Factor, BF) and designates the future active X. The A aggregate marks the X where *Xist* transcription is enhanced and is interpreted as an activating factor (AF). Importantly, the thresholds predicted by our theory for the symmetry breaking mechanisms also fall in the correct biochemical range (see above and Fig. 3 bottom panel).

The time scale required to break the symmetry in a real nucleus can depend on a number of details. Our MC provides, thus, only a very rough order of magnitude estimate. As shown in Fig. 3 top panel, such a time scale is predicted to be around 10 hours, a value of the order of the time required for XCI initiation.
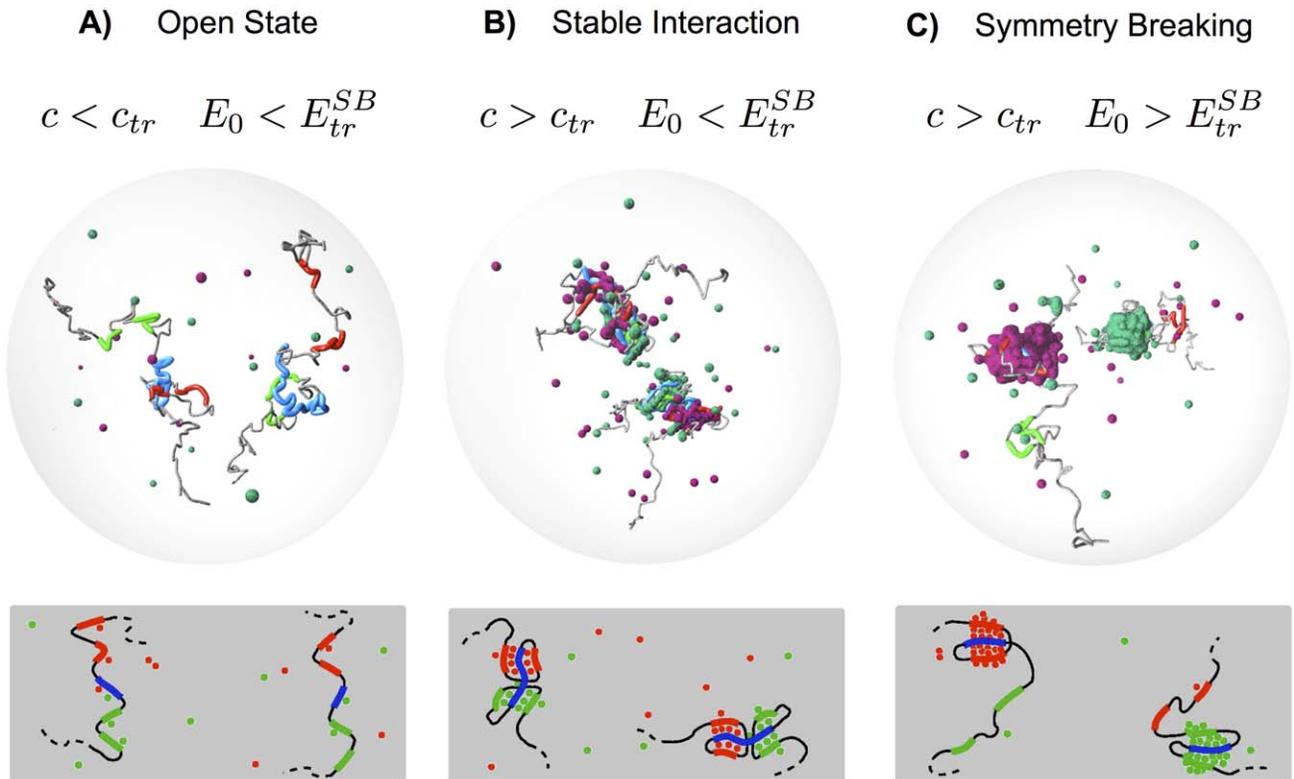
In males other processes could intervene, yet it is easy to see how the same two factors mechanism can work, i.e., why the only X is usually bound by the B aggregate (and not by A) to repress *Xist*. In fact, the affinities of A and B molecules for the type-γ region are expected, in general, to be different: $E_{X_A} \neq E_{X_B}$. Hence, if E_{X_B} is larger than E_{X_A}, it is thermodynamically convenient that B molecules bind the X, a difference of a few units in kT being sufficient to skew of orders of magnitudes the binding probability of A and B.

Finally, variants of the model can be considered to account for further biological details. For instance, additional molecular factors, or the effects on polymer colocalization can be discussed (see Text S1), but no relevant changes to the present scenario are found.

Discussion

Our schematic model (Fig. 1.B) predicts that two kinds of molecular regulators, type-A and B molecules, interact with a set of specific regions along the polymers. Current 3C data [13] suggest that our type-α and type-β regions map respectively in the area 5′ and 3′ to *Xist*, while the type-γ region is in between.

We showed that in our model only three classes of stable conformational states exist (see Fig. 4 A,B,C). The system spontaneously falls in one of them, according to molecule concentration and homotypic interaction, c and E_0. State changes are regulated by a "conformation" and by a "symmetry breaking" switch, related to two distinct thermodynamic phase transitions [25]. The switches are controlled by changing c and E_0 above/below specific threshold values. Their on/off nature can explain how a sharp regulation of nuclear architecture and stochastic choice of fate can be reliably obtained by simple strategies, such as protein upregulation or chromatin modification. Importantly, the model predicts energy/concentration thresholds which are in the expected biological range (weak biochemical energies, fractions of μmole/litre concentrations).

A) Open State

$$c < c_{tr} \quad E_0 < E_{tr}^{SB}$$

B) Stable Interaction

$$c > c_{tr} \quad E_0 < E_{tr}^{SB}$$

C) Symmetry Breaking

$$c > c_{tr} \quad E_0 > E_{tr}^{SB}$$

Figure 4. System states and transitions. The figure summarizes the system possible states and how they change by action of the Conformation and the Symmetry Breaking switch (top pictures are from MC simulations, bottom ones are schematic drawings). The switches have a thermodynamic nature and are regulated by increasing, e.g., c and E_0 (i.e., molecule concentration and homotypic interaction) above precise threshold values, c_{tr} and E_{tr}^{SB}. **A)** For $c < c_{tr}$ and $E_0 < E_{tr}^{SB}$, the polymers are found in a random open state. **B)** For $c > c_{tr}$, a conformation change is activated: type-α and type-β regions stably interact with type-γ, and a two loop conformation is established symmetrically on the two polymers. **C)** If $E_0 > E_{tr}^{SB}$, a symmetry breaking occurs as the type-α loop persists on one, randomly selected, polymer (where type-β loop is released), whereas the other polymer takes the opposite conformation. This results from the self-assembling of a single major aggregate of type-A and of type-B molecules competing to bind to type-γ region.

doi:10.1371/journal.pcbi.1002229.g004

We now discuss how the present scenario can recapitulate in a unified framework important experimental results on XCI.

Xic architecture, "counting" and "choice"

Before XCI, the *Xic* conformation is found to be identical on the two X's [13]: *Tsix* and *Xite* genes are looped onto a "buffer" region; similarly, *Jpx*, *Xist* and the "buffer" form a second hub with *Xist*. Upon XCI, on the future active X, the *Jpx-Xist*-buffer hub opens while *Xite* remains in contact with *Tsix*. On the other X, instead, the *Tsix-Xite* interactions is lost whereas *Xist* and *Jpx* remain in contact.

Our model rationalizes how those elements are sharply regulated to recognize each other and to form stable interactions based on weak biochemical bonds. It can also explain how the same physical elements later at XCI spontaneously break the X symmetry. The molecular aggregate bound, in our model, to the type-β regions (which should encompass the *Tsix-Xite* area) is interpreted as a factor related to *Xist* silencing (i.e., to its Blocking Factor, BF [1,2,4]) and designates the future active X; the different aggregate bound to type-α regions, encompassing the *Jpx* area of the other X would be linked to an *Xist* activating factor (AF) [5,6,22]. The link between architectural changes and choice of fate emerges here naturally.

During XCI establishment, the inactive X undergoes further architectural reorganization [3,14,15]. The mechanistic details of those conformational changes are still not understood, but they could involve mechanisms as those illustrated here.

Other interesting models have been proposed for "counting&-choice" at XCI, but still none had focused on the *Xic* spatial organization, including our original Symmetry Breaking theory [19]. In the approach of ref. [21], each X chromosome is assumed to have an independent probability to initiate inactivation. Two competing factors exist: an X-linked XCI-activator and an XCI-inhibitor produced by autosomes. In a male XY cell the XCI-activator concentration is too low to initiate the inactivation of the only X; in female XX cells the initial XCI-activator concentration is, instead, above the threshold needed to start XCI. As soon as one X is inactivated, the XCI-activator concentration falls down to the levels found in males, and thus the other X remains active. A different model [22] poses that two types of sites are present on the X: "XCI-init" which is responsible for the initiation of inactivation of the X bearing it, and "XCI-repres" sites which inhibit the action of "XCI-init". Each active X produces molecules, say *A* molecules, which bind to some autosomal sites. If these sites are saturated, the autosomes produce a set of molecules *I*, which, with a "Symmetry Breaking" mechanism [19], self-assemble into a single molecular factor and inhibit the activity of "XCI-repres" sites on one of the two X, determining its inactivation. As the availability of the *A* signal is reduced, it is no longer sufficient to saturate the autosomal receptors, and the remaining X remains active.

The mechanisms for conformational changes we discussed here are rooted in thermodynamics and are, thus, very robust to difference in molecular details. They could apply then to all the mentioned models for "counting&choice". An interesting question concerns the applicability of those models to mammals other than mice. Important differences have emerged, for instance, between human and mice XCI [28,29]. As stated above, the mechanisms we discussed for *Xic* architecture in mice stem cells are very robust, yet data on other organisms are still too scarce to decide whether such mechanisms might apply elsewhere.

Xic deletions/insertions and XCI

The phenotype of key deletions along the *Xic* (see reviews in [1,2,4,19,22] and ref.s therein) can be explained by our model. The Δ65kb deletion [30] removes 65kb encompassing *Xite* and part of *Xist/Tsix*. In heterozygous females the deleted X is always inactivated. In males it leads to the inactivation of the only X; the shorter the deletion considered within the Δ65kb (see ΔAS, ΔAJ, ΔAV, Δ34 [31]), the smaller the fraction of ectopic X inactivations in a population.

Those deletions, in our model, map into sites where the *Xist* "blocking factor" (BF) binds (and blocks inactivation of that X): Δ65kb removes a large portion of binding sites, thus the deleted X has a strongly reduced affinity for the BF (w.r.t. the wild type X) which does not bind there; the shorter the deletion, the weaker the effect. So, in heterozygously deleted females a skewed random XCI occurs, whereas in males the only X can be inactivated. These deletions can also impact the formation of the BF itself because the involved regions possibly encode some of its components.

Heterozygous $Tsix^{\Delta CpG}$ [32] and $Xite^{\Delta L}$ [33] deletions in females also result in the inactivation of the deleted X. Their homozygous counterpart produces, though, "aberrant counting/chaotic choice", i.e., presence of two active or inactive X's in a fraction of the cell population [18]. While that cannot be easily rationalized by other models (see, e.g., [21]), in our framework it is originated simply because the BF can fail to bind at all [34].

ΔXTX is deletion including *Xist*, *Tsix* and *Xite*, which in heterozygous causes a skewed XCI, as only the Wild Type X gets inactivated [21]. In the frame of our model ΔXTX could have a double effect: on the one hand, it hinders the binding of the AF and BF to the deleted X, by removing a number of their binding sites; on the other it affects especially the BF, since it removes the *Tsix/Xite* genes which are presumably linked to some of the BF components. Thus, the overall effect will be that while the deleted X remains active (as it lacks *Xist*), the BF is depleted and the AF wins the competition for binding the Wild Type chromosome, which is then inactivated.

Transgenic insertions are also interesting [35]. One of the predictions of our model is the highly non-linear effect of deletion/insertion, due to the "switch-like" nature of the underlying thermodynamic mechanism. The insertion experiments of ref. [35] support this view: long *Xic* transgenes can cause inactivation on male ES cells only when they are present in multiple copies,

while single insertions do not have appreciable effects. The outcome of other deletions/insertions, such as *Xist*Δ*promoter* [36], *Xist*Δ1-5 [37], *Jpx* [5], *Rnf*12 [6], etc., are similarly explained (see Text S1). XCI in diploid cells with more than two X and in polyploid cells [21] can be understood as well in our scenario (see Text S1), but additional biological hypotheses are required, since key pieces of information are still missing.

In summary, we illustrated physical switch-like mechanisms establishing conformational changes and symmetry breaking in a polymer model. For clarity, we included just the required minimal ingredients, but our model can accommodate more realistic molecular details. It can be mapped into the *Xic* region of X chromosomes to explain their complex self-organization and other important aspects of random XCI, such as the deep connection between *Xic* architectural changes and *Xist* choice of fate, reconciling within a single framework a variety of experimental evidences. The on-off character of the underlying mechanisms can also explain how sharp and reliable regulation of XCI can be attained by simple strategies, such as gene upregulation or chromatin modification.

It supports a picture where random XCI could be governed by a few core molecular elements and basic physical processes. Two main groups of molecular factors are envisaged to control the process and to produce an activating and a blocking factor for *Xist*. The specific polymer regions in our model emerge as key cis-regulators which orchestrates functional contacts along the *Xic*. Experiments targeted at that area could test their role. The model also predicts threshold effects of, e.g., genetic deletions of the regulatory regions.

The precise nature of factors and sequences involved at XCI could differ from the minimal one considered here, yet the thermodynamic mechanisms we discussed are robust and independent of the specific molecular details. Similar mechanisms could be, thus, relevant to XCI and, more generally, to other nuclear processes requiring, for example, chromatin spatial reorganizations [38–40] or alternative choices [41].

Supporting Information

Text S1 Supplementary text and figures covering the following topics: additional details on the model, polymer colocalization, effects of *Xic* Deletion/Insertion experiments on XCI and XCI process in cells with more than two X's.
(PDF)

Acknowledgments

Author Contributions

References

1. Avner P, Heard E (2001) X-Chromosome Inactivation: counting, choice and initiation. Nat Rev Genet 2: 59–67.
2. Wutz A, Gribnau J (2007) X inactivation Xplained. Curr Opin Genet Dev 17: 387–393.
3. Chow J, Heard E (2009) X inactivation and the complexities of silencing a sex chromosome. Curr Opin Cell Biol 21: 359–66.
4. Lee JT (2009) Lessons from X-chromosome inactivation: long ncRNA as guides and tethers to the epigenome. Genes Dev 23: 1831–42.
5. Tian D, Sun S, Lee JT (2010) The long noncoding RNA, Jpx, is a molecular switch for X chromosome inactivation. Cell 143: 390–403.
6. Jonkers I, Barakat TS, Achame EM, Monkhorst K, Kenter A, et al. (2009) RNF12 is an X-Encoded dose-dependent activator of X chromosome inactivation. Cell 139: 999–1011.
7. Chureau C, Chantalat S, Romito A, Galvani A, Duret L, et al. (2011) Ftx is a non-coding rna which affects xist expression and chromatin structure within the x-inactivation center region. Hum Mol Genet 20: 705–18.
8. Navarro P, Pichard S, Ciaudo C, Avner P, Rougeulle C (2005) Tsix transcription across the xist gene alters chromatin conformation without affecting xist transcription: implications for x-chromosome inactivation. Genes Dev 19: 1474–84.

9. Sado T, Hoki Y, Sasaki H (2005) Tsix silences xist through modification of chromatin structure. Dev Cell 9: 159–65.

10. Navarro P, Chambers I, Karwacki-Neisius V, Chureau C, Morey C, et al. (2008) Molecular coupling of xist regulation and pluripotency. Science 321: 1693–5.

11. Donohoe ME, Silva SS, Pinter SF, Xu N, Lee JT (2009) The pluripotency factor Oct4 interacts with Ctcf and also controls X-chromosome pairing and counting. Nature 460: 128–32.

12. Navarro P, Oldfield A, Legoupi J, Festuccia N, Dubois A, et al. (2010) Molecular coupling of tsix regulation and pluripotency. Nature 468: 457–60.

13. Tsai CL, Rowntree RK, Cohen DE, Lee JT (2008) Higher order chromatin structure at the X-Inactivation center via looping DNA. Dev Biol 319: 416–25.

14. Splinter E, de Wit E, Nora EP, Klous P, van de Werken HJG, et al. (2011) The inactive X chromosome adopts a unique three-dimensional conformation that is dependent on xist rna. Genes Dev 25: 1371–83.

15. Chaumeil J, Le Baccon P, Wutz A, Heard E (2006) A novel role for xist rna in the formation of a repressive nuclear compartment into which genes are recruited when silenced. Genes Dev 20: 2223–37.

16. Pugacheva EM, Tiwari VK, Abdullaev Z, Vostrov AA, Flanagan PT, et al. (2005) Familial cases of point mutations in the XIST promoter reveal a correlation between CTCF binding and pre-emptive choices of X chromosome inactivation. Hum Mol Genet 14: 953–65.

17. Donohoe ME, Zhang LF, Xu N, Shi Y, Lee JT (2007) Identification of a Ctcf cofactor, Yy1, for the X chromosome binary switch. Mol Cell 25: 43–56.

18. Lee JT (2005) Regulation of X-chromosome counting by Tsix and Xite sequences. Science 309: 768–71.

19. Nicodemi M, Prisco A (2007) Symmetry-breaking model for X-chromosome inactivation. Phys Rev Lett 98: 108104.

20. Augui S, Filion GJ, Huart S, Nora E, Guggiari M, et al. (2007) Sensing X chromosome pairs before X Inactivation via a novel X-pairing region of the Xic. Science 318: 1632–1636.

21. Monkhorst K, Jonkers I, Rentmeester E, Grosveld F, Gribnau J (2008) X Inactivation counting and choice is a stochastic process: evidence for involvement of an X-linked activator. Cell 132: 410–421.

22. Starmer J, Magnuson T (2009) A new model for random X chromosome inactivation. Development 136: 1–10.

23. Doi M, Edwards S (1986) The Theory of Polymer Dynamics. Oxford, USA: Clarendon Press.

24. Nicodemi M, Prisco A (2009) Thermodynamic pathways to genome spatial organization in the cell nucleus. Biophys J 96: 2168–2177.

25. Chandler D (1987) Introduction to Modern Statistical Mechanics Oxford University Press.

26. Quitschke WW, Taheny MJ, Fochtmann LJ, Vostrov AA (2000) Differential effect of zinc finger deletions on the binding of CTCF to the promoter of the amyloid precursor protein gene. Nucleic Acids Res 28: 3370–3378.

27. Renda M, Baglivo I, Burgess-Beusse B, Esposito S, Fattorusso R, et al. (2007) Critical DNA binding interactions of the insulator protein CTCF: a small number of zinc fingers mediate strong binding, and a single finger-DNA interaction controls binding at imprinted loci. J Biol Chem 282: 33336–45.

28. Okamoto I, Patrat C, Thépot D, Peynot N, Fauque P, et al. (2011) Eutherian mammals use diverse strategies to initiate x-chromosome inactivation during development. Nature 472: 370–4.

29. van den Berg IM, Galjaard RJ, Laven JSE, van Doorninck JH (2011) Xci in preimplantation mouse and human embryos: first there is remodelling…. Hum Genet 130: 203–15.

30. Clerc P, Avner P (1998) Role of the region 3′ to Xist exon 6 in the counting process of X-chromosome inactivation. Nat Genet 19: 249–53.

31. Vigneau S, Augui S, Navarro P, Avner P, Clerc P (2006) An essential role for the DXPas34 tandem repeat and Tsix transcription in the counting process of X chromosome inactivation. Proc Natl Acad Sci U S A 103: 7390–5.

32. Lee JT, Lu N (1999) Targeted mutagenesis of Tsix leads to nonrandom X inactivation. Cell 99: 47–57.

33. Ogawa Y, Lee JT (2003) Xite, X-Inactivation intergenic transcription elements that regulate the probability of choice. Mol Cell 11: 731–43.

34. Nicodemi M, Prisco A (2007) Self-assembly and DNA binding of the blocking factor in X chromosome inactivation. PLoS Comput Biol 3: e210.

35. Heard E, Mongelard F, Arnaud D, Avner P (1999) Xist yeast artificial chromosome transgenes function as X-inactivation centers only in multicopy arrays and not as single copies. Mol Cell Biol 19: 3156–3166.

36. Penny GD, Kay GF, Sheardown SA, Rastan S, Brockdorff N (1996) Requirement for xist in x chromosome inactivation. Nature 379: 131–7.

37. Marahrens Y, Panning B, Dausman J, Strauss W, Jaenisch R (1997) Xist-deficient mice are defective in dosage compensation but not spermatogenesis. Genes Dev 11: 156–66.

38. Lanctôt C, Cheutin T, Cremer M, Cavalli G, Cremer T (2007) Dynamic genome architecture in the nuclear space: regulation of gene expression in three dimensions. Nat Rev Genet 8: 104–15.

39. Misteli T (2007) Beyond the sequence: cellular organization of genome function. Cell 128: 787–800.

40. Fraser P, Bickmore W (2007) Nuclear organization of the genome and the potential for gene regulation. Nature 447: 413–417.

41. Gimelbrant A, Hutchinson JN, Thompson BR, Chess A (2007) Widespread monoallelic expression on human autosomes. Science 318: 1136–40.