Polymer models of the organization of chromosomes in the nucleus of cells

Andrea Chiariello[1], Simona Bianco[1], Andrea Piccolo[1], Carlo Annunziatella[1], Mariano Barbieri[2], Ana Pombo[2] and Mario Nicodemi[1,*]

Presentata dal socio Leonardo Merola

(Adunanza del 19 dicembre 2014)

Keywords: genome spatial organization, polymer physics, computer simulations.

Abstract - Understanding the mechanisms that control the organization of chromosomes in the space of the nucleus of cells, and its contribution to gene regulation, is a key open issue in molecular biology. New technologies have shown that chromosomes have a complex 3D organization, which dynamically changes across organisms and cell types. To understand such a fascinating complexity, quantitative models from polymer physics have been developed, to find the principles of chromosome folding, their origin and function. Here, we provide a short review of recent progress in such an important research field where Physical and Life Sciences meet.

Riassunto - La comprensione dei meccanismi che controllano l'organizzazione dei cromosomi nello spazio del nucleo cellulare, e il loro contributo alla regolazione genica, è una questione centrale e di grande attualità in biologia molecolare. Nuove tecnologie hanno mostrato che i cromosomi hanno una complessa organizzazione 3D, diversa in organismi e tipi cellulari diversi. Per comprendere la complessità di questi fenomeni, sono stati sviluppati modelli quantitativi di fisica dei polimeri, al fine di capire i principi dell'organizzazione cromosomica, la sua origine e funzione. Questa breve nota riassume alcuni dei recenti progressi di questo importante campo di ricerca dove la fisica e le scienze della vita si incontrano.

---

[1]*Universita' di Napoli "Federico II", Dipartimento di Fisica, INFN Sezione di Napoli, CNR-SPIN, Complesso Universitario di Monte S. Angelo, Via Cintia, 80126 Napoli, Italy.*
E-mail: mario.nicodemi@na.infn.it
[2]*Berlin Institute for Medical Systems Biology, Max Delbrück Center for Molecular Medicine, Robert Rössle Strasse 10, 13125 Berlin-Buch, Germany.*
E-mail: ana.pombo@mdc-berlin.de
[*]**To whom correspondence should be addressed**

1 – Introduction: chromosome spatial organization

In the cell nucleus of higher organisms, such as mammals, chromosomes have a complex, far from random spatial organization (see, e.g., [1-3]). A well-described example is the co-localization of the *Xist* locus on the two X chromosomes in female mammalian cells during X-inactivation, but a number of other cases have been reported. Chromosome organization serves vital functional purposes, disruptions being linked to a variety of diseases, including cancer (see, e.g., [1-3] and references therein). New technologies, such as 'chromosome-conformation-capture' based methods [3], are opening the way to probe the folding state of chromatin at a genomic level. Along with previous techniques such as FISH, these approaches are providing the first detailed exploration of the maps of contact in the cell nucleus.

Chromosome 3D organization appears to extend across spatial scales (Fig. 1). They occupy separate, yet interacting chromosomal territories [4, 1] where long-range chromatin interactions are functionally important. Each chromosome is partitioned in megabasepair-long domains enriched for internal contacts, known as topologically associated domains (TADs) [1, 5, 6], while long stretches of chromatin interact with the nuclear lamina (defining 'lamina-associated domains', LADs) [3], and with a variety of other functional compartments, such as the nucleolus [1]. A spectrum of molecular factors mediates chromosomal interactions. For instance, Transcription Factories are 50nm-wide clusters of about half a dozen Polymerases that promote proximity between different, co-transcribed genes and transcription units [7-10]. Splicing factors, the machinery that splices nascent transcripts into messenger RNA, accumulate in splicing speckles, which are often associated with active genes, while repressed chromatin associates with heterochromatic regions or Polycomb bodies [1]. Early replication origins are also clustered in replication factories, which stably reassemble in consecutive cell cycles [11].

While the understanding of chromatin folding has inspired models for many years [12-15], the increasing level of details exposed by recent experimental advances and the complexity of the observed patterns has renewed the need to develop quantitative models, especially from polymer physics. Here we review a few recent developments in such an important research field at the rapidly growing frontier between Physics and Life Sciences.

2 - Polymer models of chromosomes

The standard model of polymer physics describing a free polymer, i.e., a polymer experiencing only self-avoidance effects, is the Self-Avoiding Walk (SAW). A SAW polymer folds spontaneously in a random, dynamically changing conformation (Fig. 1). That causes a full intermingling of polymers in a mixture by entropic forces. Yet, in case polymers also experience a strong self-attraction force, they can produce discrete chromosome territories [16].

Under different conditions, entropy might also favor territorial separation between different chromosomes: models of free polymers folded into clustered loops can hardly find the space to penetrate into each other due to an effective entropic repulsion [17-19, 14]. These models help bridge the initial picture of chromosome territories fully separated by channels devoid of DNA [20, 1], and the discovery of their intermingling [4, 21]. Entropy, however, cannot be the only force behind chromosome organization as it fails to explain the variety of specific, functional contacts (e.g., enhancer-promoter interactions), and the domain structure of chromosomes (e.g. LADs, TADs).

This view has been reinvigorated by the development of 'chromosome conformation capture' (**3C**) technologies [22,23,24], such as Hi-C [25, 22, 26], which have provided the first semi-quantitative measures of chromosomal interactions at a genomic scale. Hi-C contact matrices have shown, for instance, that the average interaction probability between pairs of genomic loci decreases with their genomic separation, approximately with power-law decay in the 0.5-7Mb range [22].

The contact probability and its power law exponent, α, have been shown to be different in different organisms, cell types and chromosomes [23]. For instance, human embryonic stem cells (ESCs) have $\alpha \sim 1.6$ [23], human interphase lymphoblastoid cells have $\alpha \sim 1.1$ [22], while in metaphase HeLa cells $\alpha \sim 0.5$ [27]. The exponent α reported in different species also varies widely: in yeast $\alpha \sim 1.5$ [29], in Drosophila $\alpha \sim 0.7\text{-}0.8$ [28], in mouse ESCs the *Xist* locus has $\alpha \sim 0.7\text{-}0.9$ [6, 30]. Even in a given cell system different chromosomes can have different exponents [31, 23]. For instance, in human lymphoblastoid cells, the gene-poor chromosome 18 has α around 1.08, while the gene-dense chromosome 19 has α close to 1.3 [23]. These experiments have contradicted the idea that a single universal architecture, originally envisaged in the Fractal Globule (FG) state [22], might recapitulate chromosome architecture. The FG state [22] is a transient, highly unstable [32] conformation traversed, for example, by a free polymer

expanding from a highly confined, non-entangled state. The FG has $\alpha=1$ and also fails to describe the formation of TADs or LADs, as well as microscopy data of inter-locus distances [23].

An alternative scenario is proposed with the Strings & Binders Switch (SBS) model (Fig. 2a) [33, 34, 23]: in this view chromatin folding derives from interactions with other nuclear elements, such as molecular complexes that promote looping (e.g. contacts between co-expressed genes at transcription or replication factories) and associations with nuclear landmarks, such as the lamina.

In the simplest version of the SBS model, a chromatin filament is modeled as a "string", described by the golden standard of polymer physics, the SAW. Specific binding sites across the string have affinity to interact with diffusing binders that can crosslink distant binding sites forming loops. The SBS model can also be expanded to consider many polymers, different binder species, and interactions with nuclear landmarks [35, 36]. In the model, the position of the binding sites along the polymer reflects biological information, while the concentration and affinity of binders can change under varying conditions [23], as resulting, for instance, from the up-regulation of a corresponding gene or from chemical modifications of a chromosome sequence.

The SBS model, under simple and general settings, has been used to illustrate universal aspects of chromatin folding in quantitative terms, as deriving from polymer scaling properties: **(a)** three major folding classes exist (open, closed and Θ-point; Fig. 2b), corresponding to stable emergent phases of polymer physics (the Fractal Globule state is one of many possible transient states of the SBS model); **(b)** conformational changes can be sharply controlled via thermodynamic phase transitions, by simple strategies, e.g., protein up-regulation or chromatin modifications, which transduce, e.g., (*analog*) transcription factor levels into (*digital*) conformational switches; **(c)** that randomly diffusing binding molecules can establish and dynamically change, by thermodynamics mechanisms, architectural patterns, such as territory formation, TADs or LADs, and the looping out of large stretches of chromatin from territories (Fig. 2c); **(d)** that population and single cell microscopy and Hi-C data, such as contact probabilities (Fig. 3) and spatial distances, can be rationalized in a single framework [23]. The results of the SBS model are confirmed by similar findings in related, simplified models, such as the Dynamic Loop (DL) model [37, 14].

The picture of chromatin depicted by the SBS model is that chromosomes are a complex mixture of differently folded regions according to local specific factors, which can self-organize across

spatial scales by general physical mechanisms. A combination of single molecular factors can result in specificity of binding at different loci or domains, along with other molecular mechanisms, such as supercoiling [30] and plectoneme formation [38].

3 – Discussion

Models such as the SBS, informed with biological specificities, can also be employed to study specific genomic loci (e.g., the *Xist* locus) [36] to discriminate different biological scenarios and thus helping to identify the molecular determinants of chromatin folding. Such approaches could deliver very important applications to cases of medical relevance as they have the potential to predict chromatin interaction sites that best explain available data. Polymer models also have the potential to enable reconstructions of the 3D shape of specific chromosomes to enhance our understanding of gene regulation and other nuclear processes [39,40].

In summary, polymer physics, informed with and checked against experimental data, is substantially contributing to depict the first quantitative picture of the molecular mechanisms shaping chromosome folding. We expect that in the near future, further developments in modeling, experimental progresses and more detailed biological information, could push even further our comprehension of chromosome 3D architecture, hopefully advancing also the diagnostic, treatment and management of genomic diseases.

References

[1] Lanctot C, Cheutin T, Cremer M, Cavalli G, Cremer T: *Dynamic genome architecture in the nuclear space: Regulation of gene expression in three dimensions. Nature Rev Genet* 2007, 8:104-115.

[2] Misteli T: *Beyond the sequence: Cellular organization of genome function. Cell* 2007, 128:787-800.

[3] Bickmore WA, van Steensel B: *Genome architecture: Domain organization of interphase chromosomes. Cell* 2013, 152:1270-1284.

[4] Branco MR, Pombo A: *Intermingling of chromosome territories in interphase suggests role in translocations and transcription-dependent associations. PLoS Biol* 2006, 4:e138.

[5] Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, Hu M, Liu JS, Ren B: *Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature* 2012, 485:376-380.

[6] Nora EP, Lajoie BR, Schulz EG, Giorgetti L, Okamoto I, Servant N, Piolot T, van Berkum NL, Meisig J, Sedat J*, et al.*: *Spatial partitioning of the regulatory landscape of the X-inactivation centre. Nature* 2012, 485:381-385.

[7] Pombo A, Jackson DA, Hollinshead M, Wang Z, Roeder RG, Cook PR: *Regional specialization in human nuclei: Visualization of discrete sites of transcription by RNA polymerase III. EMBO J* 1999, 18:2241-2253.

[8] Osborne CS, Chakalova L, Mitchell JA, Horton A, Wood AL, Bolland DJ, Corcoran AE, Fraser P: *Myc dynamically and preferentially relocates to a transcription factory occupied by igh. PLoS Biol* 2007, 5:e192.

[9] Ferrai C, Xie SQ, Luraghi P, Munari D, Ramirez F, Branco MR, Pombo A, Crippa MP: *Poised transcription factories prime silent upa gene prior to activation. PLoS Biol* 2010, 8:e1000270.

[10] Schoenfelder S, Sexton T, Chakalova L, Cope NF, Horton A, Andrews S, Kurukuti S, Mitchell JA, Umlauf D, Dimitrova DS*, et al.*: *Preferential associations between co-regulated genes reveal a transcriptional interactome in erythroid cells. Nat Genet* 2010, 42:53-61.

[11] Jackson DA, Pombo A: *Replicon clusters are stable units of chromosome structure: Evidence that nuclear organization contributes to the efficient activation and propagation of s phase in human cells. J Cell Biol* 1998, 140:1285-1295.

[12] Langowski J: *Polymer chain models of DNA and chromatin. Eur Phys J E Soft Matter* 2006, 19:241-249.

[13] Emanuel M, Radja NH, Henriksson A, Schiessel H: *The physics behind the larger scale organization of DNA in eukaryotes. Phys Biol* 2009, 6:025008.

[14] Tark-Dame M, van Driel R, Heermann DW: *Chromatin folding--from biology to polymer models and back. J Cell Sci* 2011, 124:839-845.

[15] Barbieri M, Scialdone A, Gamba A, Pombo A, Nicodemi M: *Polymer physics, scaling and heterogeneity in the spatial organisation of chromosomes in the cell nucleus. Soft Matter* 2013, 9:8631-8635.

[16] Kreth G, Finsterle J, von Hase J, Cremer M, Cremer C: *Radial arrangement of chromosome territories in human cell nuclei: A computer model approach based on gene density indicates a probabilistic global positioning code. Biophys J* 2004, 86:2803-2812.

[17] Marenduzzo D, Micheletti C, Cook PR: *Entropy-driven genome organization. Biophys J* 2006, 90:3712-3721.

[18] Bohn M, Heermann DW, van Driel R: *Random loop model for long polymers. Physical review. E, Statistical, nonlinear, and soft matter physics* 2007, 76:051805.

[19] Rosa A, Everaers R: *Structure and dynamics of interphase chromosomes. Plos Comp Biol* 2008, 4:e1000153.

[20] Cremer T, Kreth G, Koester H, Fink RH, Heintzmann R, Cremer M, Solovei I, Zink D, Cremer C: *Chromosome territories, interchromatin domain compartment, and nuclear matrix: An integrated view of the functional nuclear architecture. Crit Rev Eukaryot Gene Expr* 2000, 10:179-212.

[21] Branco MR, Pombo A: *Chromosome organization: New facts, new models. Trends Cell Biol* 2007, 17:127-134.

[22] Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO*, et al.*: *Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science* 2009, 326:289-293.

[23] Barbieri M, Chotalia M, Fraser J, Lavitas LM, Dostie J, Pombo A, Nicodemi M: *Complexity of chromatin folding is captured by the strings and binders switch model. PNAS* 2012, 109:16173-16178.

[24] Dekker J, Rippe K, Dekker M, Kleckner N: *Capturing chromosome conformation. Science* 2002, 295:1306-1311.

[25] Tolhuis B, Palstra RJ, Splinter E, Grosveld F, de Laat W: *Looping and interaction between hypersensitive sites in the active beta-globin locus. Mol Cell* 2002, 10:1453-1465.

[26] Rodley CD, Bertels F, Jones B, O'Sullivan JM: *Global identification of yeast chromosome interactions using genome conformation capture. Fungal Genet Biol* 2009, 46:879-886.

[27] Naumova N, Imakaev M, Fudenberg G, Zhan Y, Lajoie BR, Mirny LA, Dekker J: *Organization of the mitotic chromosome. Science* 2013, 342:948-953.

[28] Sexton T, Yaffe E, Kenigsberg E, Bantignies F, Leblanc B, Hoichman M, Parrinello H, Tanay A, Cavalli G: *Three-dimensional folding and functional organization principles of the Drosophila genome. Cell* 2012, 148:458-472.

[29] Duan Z, Andronescu M, Schutz K, McIlwain S, Kim YJ, Lee C, Shendure J, Fields S, Blau CA, Noble WS: *A three-dimensional model of the yeast genome. Nature* 2010, 465:363-367.

[30] Benedetti F, Dorier J, Burnier Y, Stasiak A: *Models that include supercoiling of topological domains reproduce several known features of interphase chromosomes. Nucleic Acids Res* 2013.

[31] Kalhor R, Tjong H, Jayathilaka N, Alber F, Chen L: *Genome architectures revealed by tethered chromosome conformation capture and population-based modeling. Nature Biotech* 2011, 30:90-98.

[32] Schram RD, Barkema GT, Schiessel H: *On the stability of fractal globules. J Chem Phys* 2013, 138:224901.

[33] Nicodemi M, Panning B, Prisco A: *A thermodynamic switch for chromosome colocalization. Genetics* 2008, 179:717-721.

[34] Nicodemi M, Prisco A: *Thermodynamic pathways to genome spatial organization in the cell nucleus. Biophys J* 2009, 96:2168-2177.

[35] Scialdone A, Nicodemi M: *Diffusion-based DNA target colocalization by thermodynamic mechanisms. Development* 2010, 137:3877-3885.

[36] Scialdone A, Cataudella I, Barbieri M, Prisco A, Nicodemi M: *Conformation regulation of the X chromosome inactivation center: A model. PLoS Comput Biol* 2011, 7:e1002229.

[37] Bohn M, Heermann DW: *Diffusion-driven looping provides a consistent framework for chromatin organization. PloS One* 2010, 5:e12218.

[38] Le TB, Imakaev MV, Mirny LA, Laub MT: *High-resolution mapping of the spatial organization of a bacterial chromosome. Science* 2013, 342:731-734.

[39] Umbarger MA, Toro E, Wright MA, Porreca GJ, Bau D, Hong SH, Fero MJ, Zhu LJ, Marti-Renom MA, McAdams HH*, et al.*: *The three-dimensional architecture of a bacterial genome and its alteration by genetic perturbation. Mol Cell* 2011, 44:252-264.

[40] Nagano T, Lubling Y, Stevens TJ, Schoenfelder S, Yaffe E, Dean W, Laue ED, Tanay A, Fraser P: *Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. Nature* 2013, 502:59-64.

[41] Parada LA, McQueen PG, Misteli T: *Tissue-specific spatial organization of genomes. Genome Biol* 2004, 5:R44.

Figures

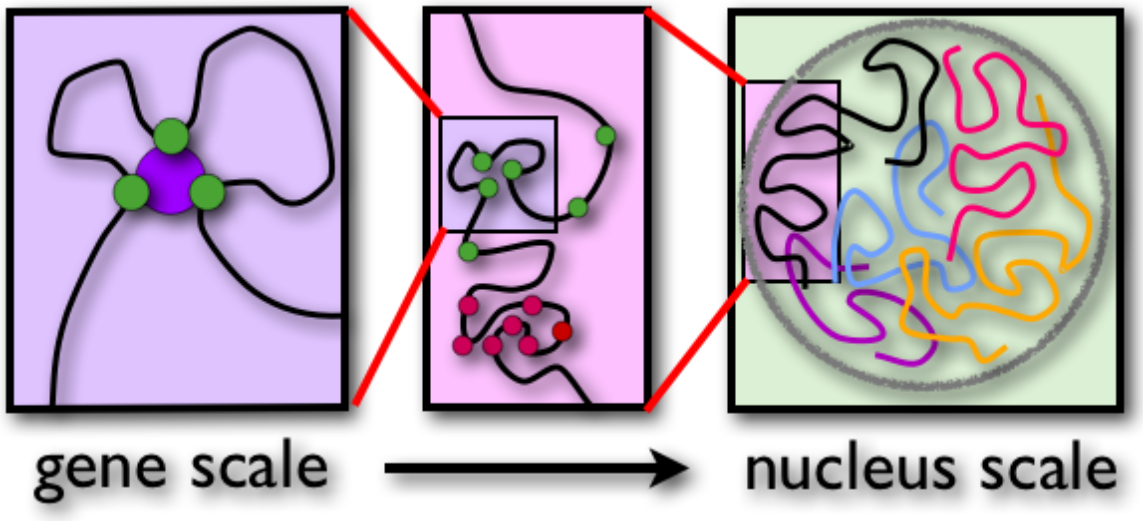

Fig. 1. Chromosome organization at different scales

Chromosomes have different levels of organization in space in the cell nucleus. Examples (from left to right): co-expressed gene co-localization at Transcription Factories; Topological Associated Domains (i.e., DNA domains with increased levels of intra- associations); chromosomal territories.

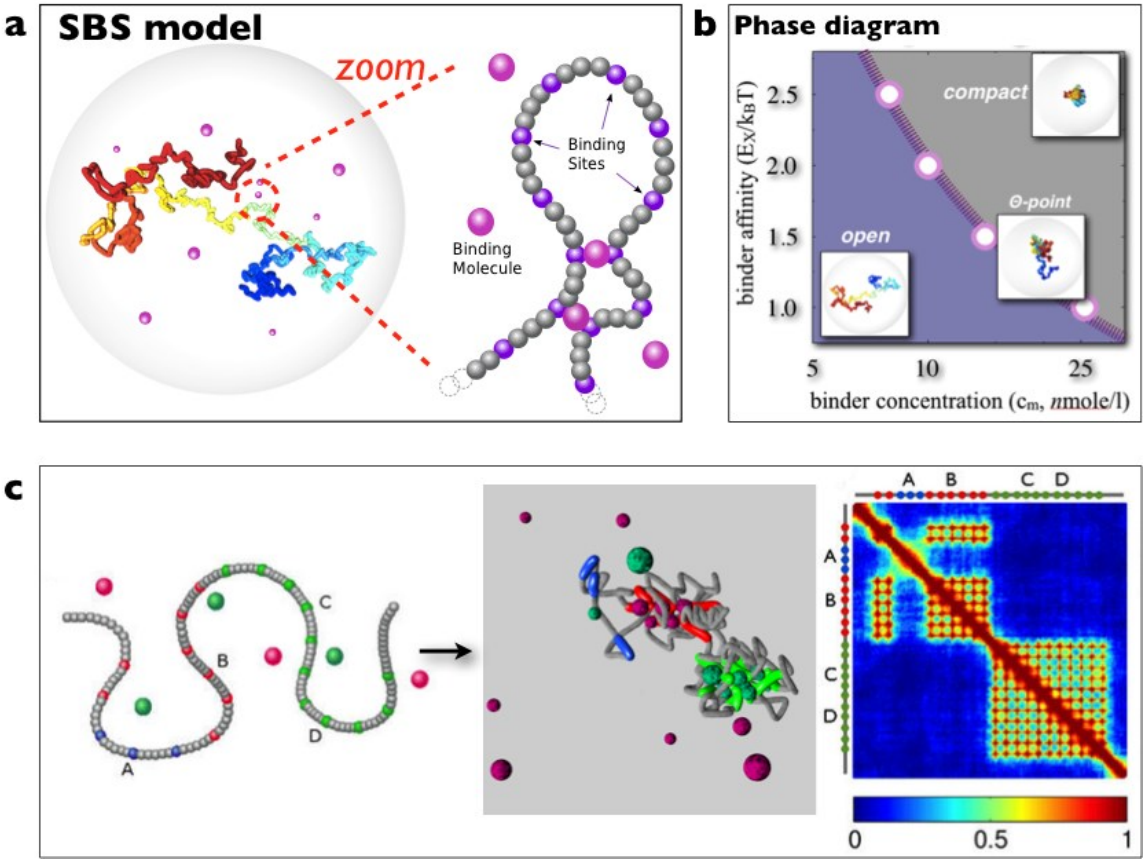

Fig. 2. The Strings & Binders Switch (SBS) model

a) In the SBS model [33, 23], chromosomes are represented by self-avoiding-walk (SAW) strings. They have binding sites for diffusing binders, which can produce loops. The SBS model parameters are binder concentrations, c_m, and their affinity for the polymer binding sites, E_X.

b) The model thermodynamic phases correspond to its different, stable conformational classes: there is a phase where the polymer folds in a random open conformation (in the universality class of the free SAW) and a phase where it spontaneously folds into a compact closed conformation. At the phase transition point, there is the Θ-point state. Conformational modifications can be controlled switch-like by crossing the phase boundary through changes in binder affinity or concentration, with no need of parameter fine-tuning.

c) Within the SBS model, the formation of chromatin domains and looping naturally emerge by the specialization of binding sites and binders. The corresponding contact matrices are close to those found experimentally by Hi-C methods.

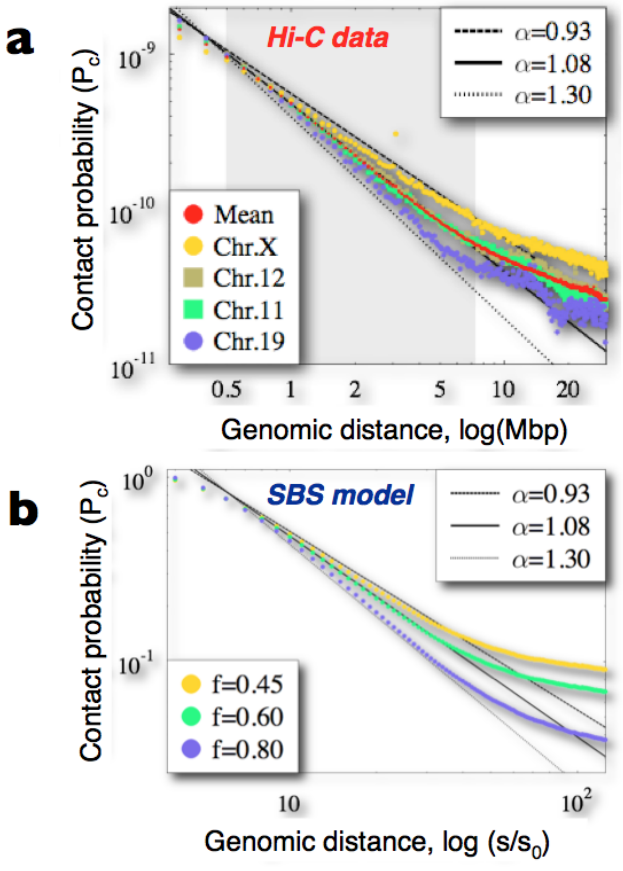

Fig. 3. Contact probabilities predicted by the SBS model match those from Hi-C.

a) The average contact probability, $P_c(s)$, as recorded in Hi-C experiments, measures the extent of chromatin interactions between pairs of loci separated by a given genomic distance. $P_c(s)$, is found to decrease with the genomic separation, s, with a power law decay at least within the 0.5-7Mb range, $P_c(s) \sim 1/s^{\alpha}$ [22]. The exponent, $\boldsymbol{\alpha}$, of the power law is different in different cell types, organisms and chromosomes [23].

b) In the SBS model, $\boldsymbol{\alpha}$ is ***2.1*** in the open state, ***1.5*** in the Θ-point state, and ***0*** in the compact state. The range of exponents found experimentally can be explained by a mixture of polymer states with a fraction, f, of open and a fraction, $1-f$, of compact regions as f varies. Within the SBS also the systematic bending of $P_c(s)$ above the power-law behaviour at larger genomic separations is explained.