



## UvA-DARE (Digital Academic Repository)

### Chromatin folding - from biology to polymer models and back

Tark-Dame, M.; van Driel, R.; Heermann, D.W.

**DOI**

[10.1242/jcs.077628](https://doi.org/10.1242/jcs.077628)

**Publication date**

2011

**Document Version**

Final published version

**Published in**

Journal of Cell Science

[Link to publication](#)

**Citation for published version (APA):**

Tark-Dame, M., van Driel, R., & Heermann, D. W. (2011). Chromatin folding - from biology to polymer models and back. *Journal of Cell Science*, 124(6), 839-845. <https://doi.org/10.1242/jcs.077628>

**General rights**

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

**Disclaimer/Complaints regulations**

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

# Chromatin folding – from biology to polymer models and back

Mariliis Tark-Dame<sup>1</sup>, Roel van Driel<sup>1</sup> and Dieter W. Heermann<sup>2</sup>

<sup>1</sup>Swammerdam Institute for Life Sciences, University of Amsterdam, PO Box 94215, 1090GE Amsterdam, The Netherlands

<sup>2</sup>Institute for Theoretical Physics, University of Heidelberg, Philosophenweg 19, 69120 Heidelberg, Germany

\*Author for correspondence ([heermann@tphys.uni-heidelberg.de](mailto:heermann@tphys.uni-heidelberg.de))

*Journal of Cell Science* 124, 839–845

© 2011. Published by The Company of Biologists Ltd

doi:10.1242/jcs.077628

## Summary

There is rapidly growing evidence that folding of the chromatin fibre inside the interphase nucleus has an important role in the regulation of gene expression. In particular, the formation of loops mediated by the interaction between specific regulatory elements, for instance enhancers and promoters, is crucial in gene control. Biochemical studies that were based on the chromosome conformation capture (3C) technology have confirmed that eukaryotic genomes are highly looped. Insight into the underlying principles comes from polymer models that explore the properties of the chromatin fibre inside the nucleus. Recent models indicate that chromatin looping can explain various properties of interphase chromatin, including chromatin compaction and compartmentalisation of chromosomes. Entropic effects have a key role in these models. In this Commentary, we give an overview of the recent conjunction of ideas regarding chromatin looping in the fields of biology and polymer physics. Starting from simple linear polymer models, we explain how specific folding properties emerge upon introducing loops and how this explains a variety of experimental observations. We also discuss different polymer models that describe chromatin folding and compare them to experimental data. Experimentally testing the predictions of such polymer models and their subsequent improvement on the basis of measurements provides a solid framework to begin to understand how our genome is folded and how folding relates to function.

**Key words:** Chromatin, Chromosome, Modelling, Polymer, Loops

## Introduction

Human cells contain 46 chromatin fibres, i.e. the chromosomes, with a total length of ~5 cm of nucleosomal filaments arranged like beads on a string. Packaging this in an interphase nucleus of typically 5–20 µm in diameter requires extensive folding. In the past decades, considerable evidence was accumulated showing that chromatin folding is closely related to genome function. Tightly packed and transcriptionally silent heterochromatin, and more-open transcriptionally active euchromatin represent two classic folding states. In the past decade, we started to see some first principles of chromatin folding. One is that individual chromosomes occupy discrete domains in the interphase nucleus – named chromosome territories – which intermingle only to a limited extent (Cremer and Cremer, 2010). Similarly, different parts of a chromosome also only interact very little (Dietzel et al., 1998; Goetze et al., 2007a; Goetze et al., 2007b). Another organisational principle is based on the finding that mammalian interphase chromosomes are made up of a large number of structural domains, each of which are on average ~1 Mb, that correspond to DNA replication units (Ryba et al., 2010). Furthermore, recent experimental data show that chromatin loops mediated by specific chromatin–chromatin interactions are an important aspect of chromatin organisation, because they bring together distant regulatory elements that control gene expression, such as promoters and enhancers (Carter et al., 2002; Kadauke and Blobel, 2009). Methods that are based on the chromosome conformation capture (3C) technology, which determines two distant genomic sequence elements that are in close proximity in the nucleus (Simonis et al., 2007), have revealed the presence of a large number of intra-chromosomal chromatin–chromatin interactions. The resulting loops vary in length from a few kb to up to tens of Mb and

are different in different cell types (Lieberman-Aiden et al., 2009; Simonis et al., 2006). Moreover, it has been shown for many loci that changes in transcriptional activity are tightly correlated with changes in folding (Sproul et al., 2005). Together, these observations show that packaging of the chromatin fibre in the interphase nucleus is closely related to genome function and that loops are a prominent feature of interphase chromatin.

The notion that chromatin loops are important for overall genome organisation is also supported from the perspective of polymer models. Recent polymer modelling efforts show that the formation of loops can endow polymers such as chromatin with properties that explain several of its key properties, including chromatin compaction and compartmentalisation. The importance of polymer models is that they aim to explain properties of chromatin on the basis of physical principles and discard those models that do not fulfil this criterion. They make precise predictions that can be tested experimentally and their outcome can be used to further improve the model, thereby increasing our understanding of chromatin folding.

In this Commentary, we demonstrate that efforts to unravel the complex and dynamic relationship between eukaryotic gene regulation and chromatin folding benefit from a marriage between polymer physics and cell biology. We will briefly summarise what is known about the molecular basis and functional role of chromatin folding, before discussing recent insights into chromatin folding that have been obtained from polymer models.

## Chromatin looping – linking chromatin folding to genome function

Chromatin looping is defined as the physical interaction between two sequence elements on the same chromosome. The idea that the

chromatin fibre forms loops is already several decades old, but only about 10 years ago the notion grew that looping has a direct role in gene regulation (Bulger and Groudine, 1999). However, only recently it has become possible to directly measure chromatin–chromatin interactions (i.e. looping) by mapping DNA sequences that physically interact using the 3C technology. Several recent reviews have addressed various aspects of chromatin looping and the reader is referred to these for details (Gondor and Ohlsson, 2009; Kadauke and Blobel, 2009; Sexton et al., 2009; Zlatanova and Caiafa, 2009a). Genome-wide mapping of chromatin–chromatin interactions in cultured human lymphoblasts revealed that human chromosomes form an unexpectedly large number of loops with sizes of up to tens of Mb (Lieberman-Aiden et al., 2009). There is growing evidence that distant regulatory elements, such as enhancers, physically interact with promoters and other regulatory sequences. Two well-studied examples are the complex looping of the developmentally controlled  $\beta$ -globin locus and the maternally or paternally imprinted H19-Igf2 locus (Han et al., 2008; Nativio et al., 2009; Noordermeer and de Laat, 2008). It has been shown that these intra-locus loops, which typically are in the range of one to a few tens of kb, can result in gene activation as well as inhibition.

Larger loops are formed by the formation of transcription factories, nuclear structures that contain several transcriptionally active genes (Cook, 2010; Mitchell and Fraser, 2008; Sutherland and Bickmore, 2009). In addition to the short-range interactions within the  $\beta$ -globin locus, the enhancer-like control region (LCR) within the  $\beta$ -globin locus interacts with loci that are located many Mb away (Kooren et al., 2007; Simonis and de Laat, 2008). Another example of long-range looping is the clustering of distant polycomb response elements in *Drosophila melanogaster* (Sexton et al., 2009) and of insulator elements found in higher eukaryotes (Bushey et al., 2008). Evidently, interphase chromosomes form an intricate network of loops.

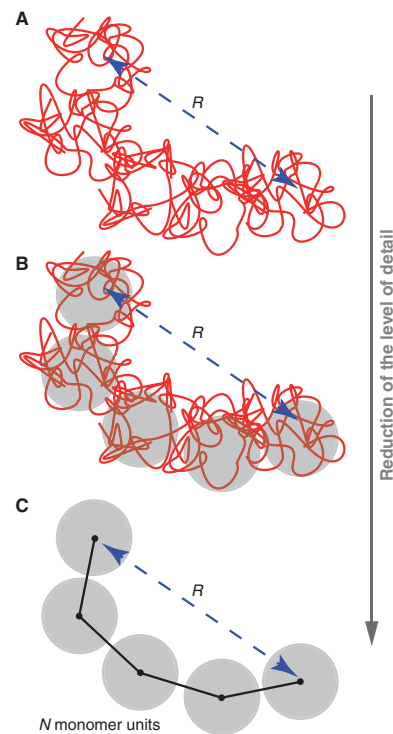
Several proteins are involved in chromatin–chromatin interactions but two proteins stand out as being particularly important in the formation of looped chromatin structures: cohesin and CCCTC binding factor (CTCF), both of which are ubiquitously and abundantly expressed (Kim et al., 2007; Phillips and Corces, 2009; Wendt et al., 2008; Zlatanova and Caiafa, 2009b). CTCF binds to an  $\sim 20$  bp consensus sequence. About 14,000 CTCF-binding sites are present in the human genome and essentially all of these are occupied by CTCF (Kim et al., 2007). CTCF is an important regulator of gene expression by physically linking distant regulatory sequences, including promoters and enhancers (Kurukuti et al., 2006; Splinter et al., 2006) as, for example, shown for the  $\beta$ -globin gene cluster, and the paternal and maternal imprinting of the IGF2-H19 locus (Han et al., 2008; Nativio et al., 2009; Noordermeer and de Laat, 2008). Chromatin looping mediated by CTCF appears to require cohesin. Chromatin immunoprecipitation (ChIP) experiments have shown that CTCF and cohesin colocalise on CTCF-binding sites (Parelho et al., 2008; Wendt et al., 2008). Cohesin was originally identified as a protein responsible for sister chromatid cohesion during mitosis, supposedly by forming a proteinaceous ring around two chromatin fibres, and is likely to mediate chromatin looping in a similar way (Carretero et al., 2010; Dorsett, 2009; Merckenschlager, 2010). Recently, it was shown that cohesin is also able to loop chromatin independently of CTCF (Kagey et al., 2010; Schmidt et al., 2010). Considering the large number of genomic binding sites for CTCF and cohesin, together they significantly contribute to the intricate chromatin-looping

pattern that has been observed by 3C measurements.

In addition, transcription factors also appear to have a role in chromatin looping, for instance of the  $\beta$ -globin locus (Drissen et al., 2004), and in the formation of transcription factories (Sutherland and Bickmore, 2009). By combining 3C-based methods with ChIP, the genome-wide chromatin-looping network that is dependent on the cell-type-specific transcription factors Kruppel-like factor 1 (Klf1) and estrogen receptor  $\alpha$  (ER- $\alpha$ ) has been charted (Fullwood et al., 2009; Schoenfelder et al., 2010). The above studies have shown that these transcription factors are necessary to bring together those genes they co-regulate, most probably by forming transcription factories.

### Polymer models for chromosomes without loops

All polymer models that address chromatin folding assume a linear unbranched polymer that represents the chromatin fibre. Necessarily, such models are coarse-grained in that they do not describe all details of the fibre (Fig. 1). Rather, they assume that polymers are made up of monomer units connected by a flexible connector (Paul et al., 1991). In practice, such polymer models assume up to 10,000 monomers (a current computational limit), resulting in a typical chromosome of 100 Mb with a minimal monomer size of  $\sim 10$  kb. This implies that all chromatin properties



**Fig. 1. Coarse-grained chromosome.** Coarse-graining is the basis of all polymer models of chromosomes. (A) Schematic view of the full chromosome with all the details. Eliminating details below the persistence length (see Box 1) of the chromosome results in a linear (i.e. unbranched) backbone. Statistically, the polymer chain retains the same large-scale conformational properties of the chromosome without any small-scale details. The small-scale properties of the chromosome contribute in an averaged way to the properties of the monomers. (B) Chromosome with effective monomers. The monomer replaces the details of the chromosome on small scales. (C) Resulting coarse-grained backbone of the chromosome that models the chromosome on scales above the persistence length.

within a 10 kb range are averaged. In the analysis of large-scale chromatin folding such an approach is justified as long as the monomer size is larger than the persistence length  $L_p$  of the chromatin fibre, i.e. the length over which the fibre is stiff (Box 1). It should be noted that the precise value of  $L_p$  for chromatin is not known.

There are two basic types of model polymer. The random walk (RW) model is characterised by the assumption that the individual monomers have no volume. This model is also called phantom chain, Gaussian chain or worm-like polymer chain (Box 1). In this type of model, two or more monomers can occupy the same spatial position at the same time. Hence, the monomers are treated as if they have no physical dimensions and thus do not occupy any space. In contrast to the RW model, the self-avoiding walk (SAW) polymer model takes the volume of the subunits into account, implying that two monomers cannot occupy the same space at the same time. The overall consequence is that – if all other parameters are kept constant – SAW polymers are more swollen than RW polymers owing to the reduced number of possible folding conformations. In their simplest forms, RW and SAW models assume that there are no attractive or repulsive forces between monomers.

### Comparing measurements to model predictions

Often, experimental observations are reconciled with model predictions by fitting the model parameters to the experimental results. A better way to relate models to experimental data is to use variables that do not depend on parameters that can be fitted to the model. Polymer models offer several such variables. One example is the dimensionless scaling exponent  $\nu$ , which describes the relationship between the distance  $N$  of two monomers along the polymer and their physical distance  $R$  in 3D space according to the equation  $\langle R^2 \rangle \sim N^{2\nu}$  that holds for all basic polymer models (Box 1). The scaling exponent  $\nu$  classifies the folding properties of the polymer, and is different for different polymer models (e.g. the RW and SAW models), therefore allowing for an objective comparison between a model and the experimental data. The fact that  $\nu$  is independent of the monomer size in the respective model underpins the idea that the monomer size is irrelevant to the overall properties of the polymer.

Using the scaling exponent  $\nu$  to relate models to experimental data can be illustrated as follows. Chromatin folding in the interphase nucleus can be measured by 3D light microscopy using fluorescent in situ hybridisation (FISH). Systematic measurements, in 3D inside the nucleus, of the physical distance  $R$  (in  $\mu\text{m}$ ) between pairs of fluorescent DNA probes that mark specific positions on the chromatin fibre at a genomic distance  $g$  (in kb or Mb) yields a value for the scaling exponent  $\nu$  according to the relationship  $\langle R^2 \rangle \sim g^{2\nu}$  (Box 1 and Fig. 2A). If chromosomes behave according to the RW or SAW model, one should observe a linear increase of the mean square spatial distance  $\langle R^2 \rangle$  as a function of the genomic distance  $g$  (as shown in Fig. 2B). Extensive measurements of the genomic distances have been performed at different scales of genomic length in primary human fibroblasts (Mateos-Langerak et al., 2009) and the results for large genomic distances (>5 Mb) do not show the predicted continuous increase of  $\langle R^2 \rangle$ . Instead, the spatial distance  $R$  reaches a plateau beyond 5–10 Mb (Fig. 2C), essentially invalidating the two most simple polymer models RW and SAW.

Recently, Emanuel and colleagues proposed a polymer model that correctly predicts the levelling off of  $\langle R^2 \rangle$  as a function of  $g$  as

### Box 1. Basic polymer parameters and models

**Persistence length ( $L_p$ ).** Quantification for the stiffness of a polymer. Below the persistence length a polymer can be considered as a stiff elastic rod. Details below this scale are thus not relevant. Estimates for the persistence length of chromatin range from 100–200 nm (Dekker et al., 2002; Langowski, 2006).

**Excluded volume.** Two objects can not occupy the same position in space at the same time. This has important consequences for the statistics of polymer conformations and, therefore, the relationship between the 3D distance  $R$  of two points on the polymer chain and the number of monomers  $N$  between them:  $\langle R^2 \rangle \sim N^{2\nu}$ .

**Chain length ( $N$ ).** The number of monomers that the polymer chain is composed of.

**Random walk (RW).** A polymer, in which excluded volume is neglected, is called a random walk polymer. For this type of polymer chain, the following relationship exists between the physical distance between the polymer end-points  $R$  and the number of monomers  $N$ :  $\langle R^2 \rangle = L_p^2 N^{2\nu}$ , with the scaling exponent of  $\nu=0.5$ .

**Scaling exponent ( $\nu$ ).** Classification for the spatial property of the chain, i.e. how the space is filled with the polymer. For a random walk polymer the scaling exponent  $\nu$  is 0.5, whereas the value for a globular polymer is 1/3. Note that the scaling exponent does not depend on details of the chain below its persistence length.

**Self-avoiding walk (SAW).** A polymer in which the excluded volume is explicitly taken into account is a self-avoiding walk polymer. For this type of polymer chain, the following relationship exists between the physical distance between the polymer end-points  $R$  and the number of monomers  $N$ :  $\langle R^2 \rangle = f(L_p) N^{2\nu}$ , whereby the scaling exponent  $\nu=0.588$ . Thus the SAW polymer is more swollen compared with the RW polymer.

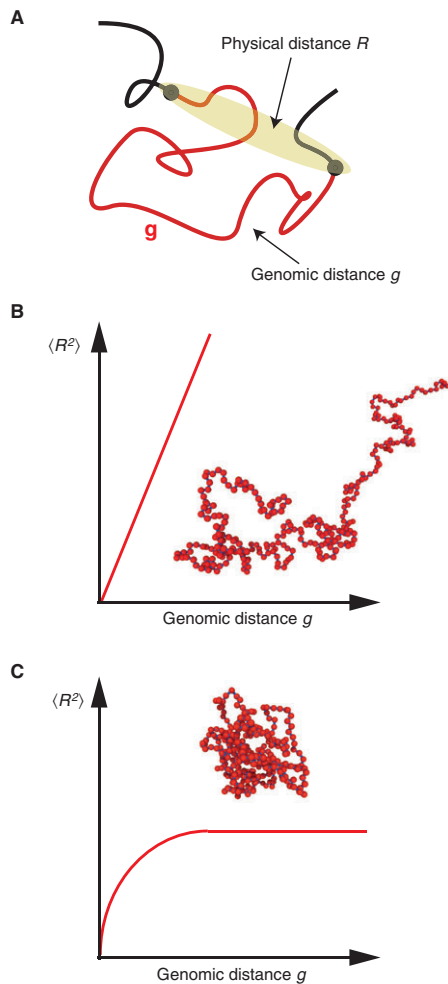
**Globular state and fractal globular state model.** A polymer is considered compact or globular if its characteristic size scales with  $\nu=1/3$ . Hence the polymer is densely packed (recall that the third root of volume of the polymer is proportional to its length). A fractal globular state model describes a knot-free polymer conformation that is packed with maximal density.

**Dynamic random loop model.** This model uses a linear backbone polymer that dynamically folds and builds loops of all length scales (Bohn et al., 2007; Mateos-Langerak et al., 2009). Here, the scaling exponent  $\nu$  becomes 0, as the length  $N$  of the polymer exceeds a certain length (2 Mb). Beyond that length, the volume occupied by the polymer remains approximately constant with increasing  $N$ .

found experimentally (Emanuel et al., 2009) (Fig. 2C). However, they include an additional constraint to the model by assuming that, through an unspecified mechanism, the volume of a folded polymer is confined to a pre-defined volume that is equivalent to a chromosome territory. This assumption forces  $\langle R^2 \rangle$  to reach a plateau value, making the outcome trivial. By contrast, we will show below that polymer models that assume chromatin looping can correctly predict the levelling off of  $\langle R^2 \rangle$  at large genomic distances without requiring additional assumptions.

### Polymer models and cell-to-cell variability in chromatin folding

Different polymer models can also be distinguished by their prediction of cell-to-cell variability of chromatin folding. A polymer model describes the ensemble of the folding configurations of the polymer and each of these conformations describes one possible geometric structure of the polymer. The configuration of the



**Fig. 2. Relationship between physical distance and genomic distance for linear and looped polymer models.** (A) Schematic representation of the physical distance  $R$  between two points on an interphase chromosome and the genomic distance  $g$ . (B) Linear polymer model that does not assume looping. This model predicts that the mean squared distance  $\langle R^2 \rangle$  increases linearly with increasing genomic distance  $g$ . (C) Model that assumes a looped chromatin chain. In contrast to the linear model, looped models predict that  $\langle R^2 \rangle$  reaches a plateau at large genomic distances – as it has been found experimentally (Mateos-Langerak et al., 2009).

polymer statistically varies over time. This translates to a cell-to-cell variability that can be measured in a population of fixed cells – because fixation creates a snapshot of every cell – each in a different chromatin folding configuration at the time of fixation. For example, the RW polymer model predicts that the physical distance  $R$  between two defined monomers of the polymer, when measured in many cells, shows a Gaussian distribution. In general, distributions are characterised by their moments, a set of parameters that uniquely characterises the distribution. For example, the first moment is the mean value of the distribution, the second moment is its width (variance). The higher moments describe other features of the distribution. Here, the ratio between the fourth moment of the distribution and the second moment squared is of interest, as it gives a dimensionless number that is independent of any parameter that can be fitted to the experimental data. Using these moments, experimentally obtained distributions – when measured with

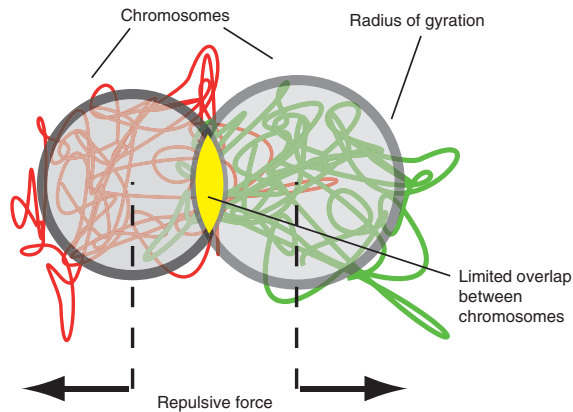
sufficiently accuracy – can be compared with polymer model predictions. For instance, a set of 3D FISH measurements obtained from a large number of individual fixed cells reflects the cell-to-cell variation and yields a distribution of the physical distance  $R$  as a function of the genomic distance (Box 1). For this distribution one can compute the moments and compare the experimental data to model predictions. Such a comparison has shown that linear polymer models, i.e. models not involving loops, are incompatible with the experimental data (Bohn and Heermann, 2009).

### Polymer models with loops

Looped polymers have a number of properties that are not observed in unlooped polymers. Entropic effects make the intermingling of two looped polymers highly unfavourable (Bohn and Heermann, 2010b), and even a small number of loops per polymer can considerably suppress mixing of polymers. Thermodynamically, such a situation can be described as repulsion between the looped polymers. This can be understood intuitively by considering a polymer that has loops covering all lengths, i.e. small, medium and large loops, relative to the total length (contour length) of the polymer. Such a polymer has a more or less sphere-like shape and is difficult to penetrate by other polymers (Fig. 3). Considering such a model predicts that chromatin loops are a key determinant of the properties of the chromatin fibre. The more loops are formed, the less space the chromatin fibre occupies and the more it is compacted. Thus, chromosomes condense as loops are formed and their intermingling is strongly reduced.

Pioneering polymer models that were developed to explain the measured properties of interphase chromatin and that take into account chromatin looping assume loops of fairly uniform size. For instance, Sachs and colleagues proposed a model in which chromatin loops of 1.5–3.5 Mb are attached to an unspecified RW backbone, with the proposed loop size being the result of fitting data on the model (Sachs et al., 1995). Others assumed that the chromatin fibre assembles in an array of rosettes of loops of uniform size (Münkel and Langowski, 1998). However, recent 3C studies of chromatin–chromatin interactions do not support the idea of loops of fixed sizes but, instead, show that chromatin loops cover a wide range of lengths, ranging from a few kb to tens of Mb (Lieberman-Aiden et al., 2009; Simonis et al., 2006). Thus far, only our dynamic random loop model (Bohn et al., 2007; Mateos-Langerak et al., 2009) and the fractal globular model developed by Dekker and co-workers (Lieberman-Aiden et al., 2009) incorporate the idea of a wide range of loop sizes.

The dynamic random loop model (Box 1) assumes a dynamic, random interaction between monomers of a polymer, creating loops that span a wide range of sizes (Bohn et al., 2007; Mateos-Langerak et al., 2009). Several characteristics of interphase chromatin folding can be explained by this model, including the observation that each interphase chromosome occupies a limited space, i.e. a chromosome territory, in the interphase nucleus. This is reflected by the levelling off of the physical distance  $R$  between pairs of sequence elements as a function of their genomic distance  $g$ ; the scaling exponent  $\nu$  becomes zero (Fig. 2C, Box 1). Furthermore, the model predicts different degrees of compaction along the length of a chromosome that are caused by variations in local looping probabilities (Bohn et al., 2007; Mateos-Langerak et al., 2009). In contrast to the dynamic random loop model, the fractal globular model (Lieberman-Aiden et al., 2009) (Box 1) is characterised by a scaling component  $\nu=1/3$  and does not predict the experimentally observed levelling off of  $R$  as a function of  $g$ . Taken together, the



**Fig. 3. Physical interaction between chromosomes.** The backbones of two chromosomes are shown in red and green. Loops and ensuing entropic effects are the major driving forces for the internal organisation of chromosome territories and their segregation in the interphase cell nucleus. Two chromosomes repel each other owing to entropic repulsion between the looped polymers. This entropic repulsion is relatively weak and, therefore, some overlap between chromosome territories occurs. The degree of overlap depends on the degree of looping of the individual chromosomes.

dynamic random loop model, which explicitly assumes looping at all lengths, therefore best explains key properties of the chromatin folding.

### Size distribution of loop size

As discussed above, dimensionless variables are parameters of choice for comparing experimental data with models. The distribution of loop sizes provides such a parameter. Using the genome-wide HiC (an extension of the 3C method) data set from the Dekker group (Lieberman-Aiden et al., 2009) allows for a quantitative analysis of the loop size distribution of chromatin in a human cell. Let us assume that two monomers of a linear polymer come in contact with each other and form a loop of the length  $l$ , where  $l$  is the number of monomers in the loop. We can then ask: what is the distribution  $p(l)$  of the lengths of the loops in the polymer? The fractal globular state model based on HiC data (Box 1) (Lieberman-Aiden et al., 2009), as well as the dynamic random loop model (Bohn and Heermann, 2010a), correctly predict a value for  $\beta$  of about 1, with  $\beta$  being the exponent in a power law that characterises the dependence of the number of loops (or their probability) as a function of the loop size. However, the fractal globular model cannot be reconciled with experimental 3D FISH data obtained from human cells (Bohn and Heermann, 2009; Mateos-Langerak et al., 2009), as they predict a cell-to-cell variation that is different from that experimentally observed, by making a wrong prediction for ratio of the higher moments for the distribution of  $R$  (Bohn and Heermann, 2009). Thus, it is apparent that, of the presently available models, the dynamic random loop polymer model most accurately describes the properties of interphase chromatin.

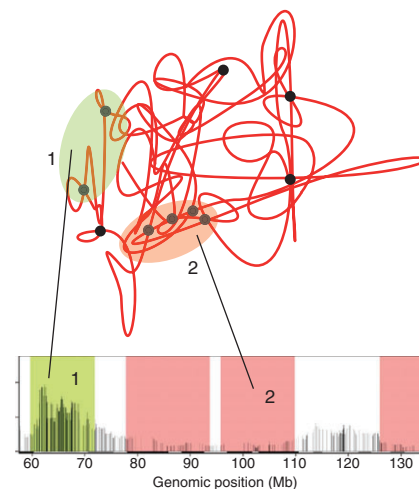
### Relationship between loops and transcription

Local chromatin folding is related to the local transcriptional activity and gene density (Goetze et al., 2007a), and to replication timing (Ryba et al., 2010). Furthermore, the transcriptional state of a chromosome and its subchromosomal domains affect chromatin

positioning inside the nucleus (Janicki et al., 2004). Transcriptionally active chromatin tends to be located nearer to the nuclear interior, whereas inactive chromatin is more frequently found closer to the nuclear periphery (Cremer et al., 2001; Dietzel et al., 2004; Goetze et al., 2007a; Goetze et al., 2007b; Scheuermann et al., 2005). Furthermore, transcriptionally active and gene-dense regions of the human genome (with a typical size of several Mb) have a more-open chromatin structure than genomic regions that are less dense in genes and less transcriptionally active (Goetze et al., 2007a) (Fig. 4). The dynamic random loop model can explain these differences by assuming a moderately higher looping density for compact chromatin domains compared with those of more-open regions (Mateos-Langerak et al., 2009). The random loop model can be further refined by using local gene expression and gene densities along chromosomes as an indicator for local loop densities. This approach should allow the model to predict modulations in chromatin compaction along the chromatin fibre (Hansjörg Jerabek and D.H., unpublished observations).

### The importance of loops for chromosome territories

Chromatin loops appear to have a dominant role in the folding of interphase chromosomes and, at the same time, they are important for the overall nuclear organisation. Cook and Marenduzzo recently investigated the effect of chromatin looping on the formation of chromosome territories, assuming that chromatin is folded in rosette-like structures with fixed loop attachment points (Cook and Marenduzzo, 2009). Their computer analyses show that such chromosomes display only limited intermingling due to entropic



**Fig. 4. Gene-expression-dependent chromatin folding.** Illustrated here is the polymer model of a looped chromosome with its linear backbone shown in red. Upon backbone folding, loops are generated that vary considerably in size. Black dots indicate looping points. The looped chromosome has an internal structure that is segregated into two types of region, both of which differ in their degree of compaction (green, open; pink, compact). The compaction of sub-chromosomal regions has been shown to depend on gene density and transcriptional activity (Goetze et al., 2007a). Below, the transcriptome map of the q-arm of human chromosome 11 is shown; vertical lines indicate individual genes, the length of the lines depicts gene activity. The genome consist of regions with high gene density and high gene activity (green), and gene deserts sparsely filled with less-active genes (red) (Caron et al., 2001; Mateos-Langerak et al., 2009).

forces and that looped chromatin fibres, indeed, repel each other (Cook and Marenduzzo, 2009).

The existence of distinct chromosome territories can also be explained with the dynamic random loop model (Bohn and Heermann, 2010a). Similarly, polymer modelling (Cook and Marenduzzo, 2009) predicts that entropic repulsion is also important at the subchromosomal level, resulting in subchromosomal domains that show little or no intermingling. This is exactly what is observed experimentally (Goetze et al., 2007a). The dynamic random loop model (Mateos-Langerak et al., 2009) suggests that chromosome segregation is driven by the formation of loops. Rosa and Everaers have proposed an alternative explanation for discrete chromosome territories and suggested that these are the consequence of the very slow rate of entanglement of chromosomes after the metaphase–interphase transition (Rosa and Everaers, 2008). They assume a SAW polymer model and do not take into account loops. All other models discussed in this overview assume that the polymer configurations are in equilibrium. Presently, there is no experimental evidence that rules out this possibility.

### Loops and the shape of chromosome territories

Another aspect of model predictions addresses the shape of interphase chromosomes. Linear polymer models, such as the RW, SAW or the globular state model (Box 1) and also the looped models, make specific predictions about the shape of chromosomes with regard to the ratios between the long and the two short axes of chromosome-equivalent ellipsoids. The shape of interphase chromosomes and chromosomal subcompartments can be experimentally analysed by FISH (Bolzer et al., 2005; Goetze et al., 2007a). Quantitative analysis of human fibroblasts shows that the deviation from a spherical chromosome shape in subchromosomal regions correlates with transcriptional activity and gene density of the chromosome; gene-rich and transcriptionally active regions appear highly non-spherical, whereas the shape of gene-poor and less active chromatin regions is more sphere-like (Goetze et al., 2007a). In addition, Khalil et al. found that the shape of chromosome territories in mouse B cells is highly non-spherical (Khalil et al., 2007).

The dynamic random loop model predicts that the elongated chromosome shape of chromosomes and chromosomal subcompartments becomes more pronounced when the loop frequency increases, offering a simple explanation for the difference in shape between gene-rich and gene-sparse chromosomal regions (Bohn and Heermann, 2010a). By contrast, globular state models do not predict strongly aspheric regions. Taken together, a looped polymer model – in addition to confirming loop distribution – can also correctly predict the experimentally observed shape of a chromosome.

### Conclusions and perspectives

In this Commentary, we show that polymer models are valuable tools in uncovering basic aspects of chromatin folding. We argue that the formation of loops has a key role in chromatin folding, owing to the entropic repulsion between looped polymers. Assuming a simple polymer backbone that forms loops of all lengths is a highly useful model for interphase chromosomes and explains many of the experimental observations. First and foremost is the correctly predicted relationship between the physical distance  $R$  between two sequence elements that can be measured by FISH, and their genomic distance  $g$ , i.e.  $\langle R^2 \rangle \sim g^{2\nu}$  (Box 1). This relationship provides a

physical basis for the understanding of chromatin folding, and establishes a link between folding and local transcriptional activity, and other biological properties. For instance, analyzing the behaviour of randomly looped polymers as a model for chromatin fibres, presents a compelling explanation why interphase chromosomes are compartmentalised and segregated into territories. The entropic repulsion between the loops and, thus, between chromosomes constitutes the physical basis of such a compartmentalisation. As discussed above, different degrees of compaction along a chromosome can also be explained with this model when assuming that looping varies along the chromosomal length. In addition, using 3C-based techniques, loops of all lengths are experimentally found, which show a characteristic size distribution that can be recapitulated with the random loop polymer model. Importantly, chromatin loops are not only key elements in chromatin folding, they are also an important component of gene regulatory systems. Hence, polymer models that incorporate looping provide a reliable framework for the analysis of the structure of chromosomes.

So far, only overall properties of chromosomes have been modelled and tested experimentally. Future work should take into account the increasing number of details that correctly predict biological function. For example, variations in local gene density and transcriptional activity along the chromosome are obvious parameters that can be incorporated into polymer models. Clearly, further experiments and modelling efforts are needed to delineate the exact relationship between genome folding and function. This requires high-volume and high-precision data sets to feed into polymer models. Another important aspect concerns the dynamics of chromatin folding and, in particular, of chromatin looping. In this aspect of future work, polymer models will be the guiding principle in designing new experimental approaches and data acquisition.

Part of this work is supported by a grant of the Foundation for Fundamental Research on Matter (FOM), which is financially supported by the Netherlands Organisation for Scientific Research (NWO) (grant FOM A29 to MTD) and of HGS MathComp Heidelberg (DWH).

### References

- Bohn, M. and Heermann, D. W. (2009). Conformational properties of compact polymers. *J. Chem. Phys.* **130**, 174901.
- Bohn, M. and Heermann, D. W. (2010a). Diffusion-driven looping provides a consistent framework for chromatin organization. *PLoS One* **5**, e12218.
- Bohn, M. and Heermann, D. W. (2010b). Topological interactions between ring polymers: Implications for chromatin loops. *J. Chem. Phys.* **132**, 044904.
- Bohn, M., Heermann, D. W. and van Driel, R. (2007). Random loop model for long polymers. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* **76**, 051805.
- Bolzer, A., Kreth, G., Solovei, I., Koehler, D., Saracoglu, K., Fauth, C., Muller, S., Eils, R., Cremer, C., Speicher, M. R. et al. (2005). Three-dimensional maps of all chromosomes in human male fibroblast nuclei and prometaphase rosettes. *PLoS Biol.* **3**, 826–842.
- Bulger, M. and Groudine, M. (1999). Looping versus linking: toward a model for long-distance gene activation. *Genes Dev.* **13**, 2465–2477.
- Bushey, A. M., Dorman, E. R. and Corces, V. G. (2008). Chromatin insulators: regulatory mechanisms and epigenetic inheritance. *Mol. Cell* **32**, 1–9.
- Caron, H., van Schaik, B., van der Mee, M., Baas, F., Riggins, G., van Sluis, P., Hermus, M. C., van Asperen, R., Boon, K., Voute, P. A. et al. (2001). The human transcriptome map: clustering of highly expressed genes in chromosomal domains. *Science* **291**, 1289–1292.
- Carretero, M., Remeseiro, S. and Losada, A. (2010). Cohesin ties up the genome. *Curr. Opin. Cell Biol.* **22**, 1–7.
- Carter, D., Chakalova, L., Osborne, C. S., Dai, Y. F. and Fraser, P. (2002). Long-range chromatin regulatory interactions in vivo. *Nat. Genet.* **32**, 623–626.
- Cook, P. R. (2010). A model for all genomes: the role of transcription factories. *J. Mol. Biol.* **395**, 1–10.
- Cook, P. R. and Marenduzzo, D. (2009). Entropic organization of interphase chromosomes. *J. Cell Biol.* **186**, 825–834.

- Cremer, M., von Hase, J., Volm, T., Brero, A., Kreth, G., Walter, J., Fischer, C., Solovei, I., Cremer, C. and Cremer, T. (2001). Non-random radial higher-order chromatin arrangements in nuclei of diploid human cells. *Chromosome Res.* **9**, 541-567.
- Cremer, T. and Cremer, M. (2010). Chromosome Territories. *Cold Spring Harb. Perspect. Biol.* **2**, a003889.
- Dekker, J., Rippe, K., Dekker, M. and Kleckner, N. (2002). Capturing chromosome conformation. *Science* **295**, 1306-1311.
- Dietzel, S., Jauch, A., Kienle, D., Qu, G. Q., Holtgreve-Grez, H., Eils, R., Munkel, C., Bittner, M., Meltzer, P. S., Trent, J. M. et al. (1998). Separate and variably shaped chromosome arm domains are disclosed by chromosome arm painting in human cell nuclei. *Chromosome Res.* **6**, 25-33.
- Dietzel, S., Zolghadr, K., Heppner, C. and Belmont, A. S. (2004). Differential large-scale chromatin compaction and intranuclear positioning of transcribed versus non-transcribed transgene arrays containing beta-globin regulatory sequences. *J. Cell Sci.* **117**, 4603-4614.
- Dorsett, D. (2009). Cohesin, gene expression and development: lessons from *Drosophila*. *Chromosome Res.* **17**, 185-200.
- Drissen, R., Palstra, R. J., Gillemans, N., Splinter, E., Grosveld, F., Philipsen, S. and de Laat, W. (2004). The active spatial organization of the  $\beta$ -globin locus requires the transcription factor EKLf. *Genes Dev.* **18**, 2485-2490.
- Emanuel, M., Radja, N. H., Henriksson, A. and Schiessel, H. (2009). The physics behind the larger scale organization of DNA in eukaryotes. *Phys. Biol.* **6**, 25008.
- Fullwood, M. J., Liu, M. H., Pan, Y. F., Liu, J., Xu, H., Mohamed, Y. B., Orlov, Y. L., Velkov, S., Ho, A., Mei, P. H. et al. (2009). An oestrogen-receptor- $\alpha$ -bound human chromatin interactome. *Nature* **462**, 58-64.
- Goetze, S., Mateos-Langerak, J., Gierman, H. J., de Leeuw, W., Giromus, O., Indemans, M. H., Koster, J., Ondrej, V., Versteeg, R. and van Driel, R. (2007a). The three-dimensional structure of human interphase chromosomes is related to the transcriptome map. *Mol. Cell Biol.* **27**, 4475-4487.
- Goetze, S., Mateos-Langerak, J. and van Driel, R. (2007b). Three-dimensional genome organization in interphase and its relation to genome function. *Semin. Cell Dev. Biol.* **18**, 707-714.
- Gondor, A. and Ohlsson, R. (2009). Chromosome crosstalk in three dimensions. *Nature* **461**, 212-217.
- Han, L., Lee, D. H. and Szabo, P. E. (2008). CTCF is the master organizer of domain-wide allele-specific chromatin at the H19/Igf2 imprinted region. *Mol. Cell Biol.* **28**, 1124-1135.
- Janicki, S. M., Tsukamoto, T., Salghetti, S. E., Tansey, W. P., Sachidanandam, R., Prasanth, K. V., Ried, T., ShavTal, Y., Berstrand, E., Singer, R. H. et al. (2004). From silencing to gene expression: real-time analysis in single cells. *Cell* **116**, 683-698.
- Kadauke, S. and Blobel, G. A. (2009). Chromatin loops in gene regulation. *Biochim. Biophys. Acta* **1789**, 17-25.
- Kagey, M. H., Newman, J. J., Bilodeau, S., Zhan, Y., Orlando, D. A., van Berkum, N. L., Ebmeier, C. C., Goossens, J., Rahl, P. B., Levine, S. S. et al. (2010). Mediator and cohesin connect gene expression and chromatin architecture. *Nature* **467**, 430-435.
- Khalil, A., Grant, J. L., Caddle, L. B., Atzema, E., Mills, K. D. and Arneodo, A. (2007). Chromosome territories have a highly nonspherical morphology and nonrandom positioning. *Chromosome Res.* **15**, 899-916.
- Kim, T. H., Abdullaev, Z. K., Smith, A. D., Ching, K. A., Loukinov, D. I., Green, R. D., Zhang, M. Q., Lobanenko, V. V. and Ren, B. (2007). Analysis of the vertebrate insulator protein CTCF-binding sites in the human genome. *Cell* **128**, 1231-1245.
- Kooren, J., Palstra, R. J., Klous, P., Splinter, E., von Lindern, M., Grosveld, F. and de Laat, W. (2007).  $\beta$ -globin active chromatin Hub formation in differentiating erythroid cells and in p45 NF-E2 knock-out mice. *J. Biol. Chem.* **282**, 16544-16552.
- Kurukuti, S., Tiwari, V. K., Tavosoidana, G., Pugacheva, E., Murrell, A., Zhao, Z. H., Lobanenko, V., Reik, W. and Ohlsson, R. (2006). CTCF binding at the H19 imprinting control region mediates maternally inherited higher-order chromatin conformation to restrict enhancer access to Igf2. *Proc. Nat. Acad. Sci. USA* **103**, 10684-10689.
- Langowski, J. (2006). Polymer chain models of DNA and chromatin. *Eur. Phys. J. E Soft Matter* **19**, 241-249.
- Lieberman-Aiden, E., van Berkum, N. L., Williams, L., Imakaev, M., Ragozy, T., Telling, A., Amit, I., Lajoie, B. R., Sabo, P. J., Dorschner, M. O. et al. (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326**, 289-293.
- Mateos-Langerak, J., Bohn, M., de Leeuw, W., Giromus, O., Manders, E. M., Verschure, P. J., Indemans, M. H., Gierman, H. J., Heermann, D. W., van Driel, R. et al. (2009). Spatially confined folding of chromatin in the interphase nucleus. *Proc. Natl. Acad. Sci. USA* **106**, 3812-3817.
- Merkenschlager, M. (2010). Cohesin: a global player in chromosome biology with local ties to gene regulation. *Curr. Opin. Genet. Dev.* **20**, 1-7.
- Mitchell, J. A. and Fraser, P. (2008). Transcription factories are nuclear subcompartments that remain in the absence of transcription. *Genes Dev.* **22**, 20-25.
- Münkel, C. and Langowski, J. (1998). Chromosome structure predicted by a polymer model. *Phys. Rev. E* **57**, 5888-5896.
- Nativio, R., Wendt, K. S., Ito, Y., Huddleston, J. E., Uribe-Lewis, S., Woodfine, K., Krueger, C., Reik, W., Peters, J. M. and Murrell, A. (2009). Cohesin is required for higher-order chromatin conformation at the imprinted IGF2-H19 locus. *PLoS Genet.* **5**, e1000739.
- Noordermeer, D. and de Laat, W. (2008). Joining the loops:  $\beta$ -globin gene regulation. *IUBMB Life* **60**, 824-833.
- Parelho, V., Hadjur, S., Spivakov, M., Leleu, M., Sauer, S., Gregson, H. C., Jarmuz, A., Canzonetta, C., Webster, Z., Nesterova, T. et al. (2008). Cohesins functionally associate with CTCF on mammalian chromosome arms. *Cell* **132**, 422-433.
- Paul, W., Binder, K., Kremer, K. and Heermann, D. W. (1991). Structure property correlation of polymers, a Monte Carlo approach. *Macromolecules* **24**, 6332-6334.
- Phillips, J. E. and Corces, V. G. (2009). CTCF: master weaver of the genome. *Cell* **137**, 1194-1211.
- Rosa, A. and Everaers, R. (2008). Structure and dynamics of interphase chromosomes. *PLoS Comp. Biol.* **4**, e1000153.
- Ryba, T., Hiratani, I., Lu, J., Itoh, M., Kulik, M., Zhang, J., Schulz, T. C., Robins, A. J., Dalton, S. and Gilbert, D. M. (2010). Evolutionarily conserved replication timing profiles predict long-range chromatin interactions and distinguish closely related cell types. *Genome Res.* **20**, 761-770.
- Sachs, R. K., van den Engh, G., Trask, B., Yokota, H. and Hearst, J. E. (1995). A random-walk/giant-loop model for interphase chromosomes. *Proc. Natl. Acad. Sci. USA* **92**, 2710-2714.
- Scheuermann, M. O., Murmann, A. E., Richter, K., Görisch, S. M., Herrmann, H., Lichter, P. (2005). Characterization of nuclear compartments identified by ectopic markers in mammalian cells with distinctly different karyotype. *Chromosoma* **114**, 39-53.
- Schmidt, D., Schwalie, P. C., Ross-Innes, C. S., Hurtado, A., Brown, G. D., Carroll, J. S., Flicek, P. and Odom, D. T. (2010). A CTCF-independent role for cohesin in tissue-specific transcription. *Genome Res.* **20**, 578-588.
- Schoenfelder, S., Sexton, T., Chakalova, L., Cope, N. F., Horton, A., Andrews, S., Kurukuti, S., Mitchell, J. A., Umlauf, D., Dimitrova, D. S. et al. (2010). Preferential associations between co-regulated genes reveal a transcriptional interactome in erythroid cells. *Nat. Genet.* **42**, 53-61.
- Sexton, T., Bantignies, F. and Cavalli, G. (2009). Genomic interactions: chromatin loops and gene meeting points in transcriptional regulation. *Semin. Cell Dev. Biol.* **20**, 849-855.
- Simonis, M. and de Laat, W. (2008). FISH-eyed and genome-wide views on the spatial organisation of gene expression. *Biochim. Biophys. Acta* **1783**, 2052-2060.
- Simonis, M., Klous, P., Splinter, E., Moshkin, Y., Willemsen, R., de Wit, E., van Steensel, B. and de Laat, W. (2006). Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). *Nat. Genet.* **38**, 1348-1354.
- Simonis, M., Kooren, J. and de Laat, W. (2007). An evaluation of 3C-based methods to capture DNA interactions. *Nat. Methods* **4**, 895-901.
- Splinter, E., Heath, H., Kooren, J., Palstra, R. J., Klous, P., Grosveld, F., Galjart, N. and de Laat, W. (2006). CTCF mediates long-range chromatin looping and local histone modification in the  $\beta$ -globin locus. *Genes Dev.* **20**, 2349-2354.
- Sprout, D., Gilbert, N. and Bickmore, W. A. (2005). The role of chromatin structure in regulating the expression of clustered genes. *Nat. Rev. Genet.* **6**, 775-781.
- Sutherland, H. and Bickmore, W. A. (2009). Transcription factories: gene expression in unions? *Nat. Rev. Genet.* **10**, 457-466.
- Wendt, K. S., Yoshida, K., Itoh, T., Bando, M., Koch, B., Schirghuber, E., Tsutsumi, S., Nagae, G., Ishihara, K., Mishiro, T. et al. (2008). Cohesin mediates transcriptional insulation by CCCTC-binding factor. *Nature* **451**, 796-801.
- Zlatanova, J. and Caiafa, P. (2009a). CCCTC-binding factor: to loop or to bridge. *Cell. Mol. Life Sci.* **66**, 1647-1660.
- Zlatanova, J. and Caiafa, P. (2009b). CTCF and its protein partners: divide and rule? *J. Cell Sci.* **122**, 1275-1284.