The fractal globule as a model of chromatin architecture in the cell

Leonid A. Mirny

July 26, 2010

Harvard-MIT Division of Health Sciences and Technology, and
Department of Physics, Massachusetts Institute of Technology,
Cambridge, USA

Abstract

The fractal globule is a compact polymer state that emerges during polymer condensation as a result of topological constraints that prevent one region of the chain from passing across another one. This long-lived intermediate state was introduced in 1988 (Grosberg et al., 1988) and has not been observed in experiments or simulations until recently (Lieberman-Aiden, van Berkum et al., 2009). Chromatin capture characterization of human chromatin demonstrated that the fractal globule is the only conformational ensemble of a polymer that is consistent with the data (Lieberman-Aiden, van Berkum et al., 2009).

Here, we present the concept of the fractal globule, comparing it to other states of a polymer and focusing on its properties relevant for biophysics of chromatin. We then discuss properties of the fractal globule that make it an attractive model for chromatin organization inside a cell. Next, we connect the fractal globule to recent studies that emphasize topological constraints as a primary factor driving formation of chromosomal territories. We discuss how predictions of fractal globule organization can be tested experimentally. Finally, we discuss whether fractal globule architecture can be relevant for chromatin packing in other organisms such as yeast and bacteria.

Introduction: an ensemble of structures

How are meters of DNA packed inside the 5μm-diameter nucleus of a cell? Recent developments in imaging (Berger et al., 2008; Cremer and Cremer, 2010; Müller et al., 2004; Strickfaden et al., 2010; Yokota et al., 1995) and chromosome capture techniques (Duan et al., 2010; Miele and Dekker, 2009; Ohlsson and Göndör, 2007; van Berkum and Dekker, 2009) provided new insights into this problem. Before looking at specific observations, however, it is worth asking a question: what kind of DNA structures do we expect to find in this packing?

Fifty years of research in Structural Biology have provided tens of thousands of protein and nucleic acid structures resolved to a fraction of nanometer. Such high resolution is possible because billions of copies of a particular protein or a nucleic acid all have precisely the same shape in the individual cells of a crystal. Moreover, NMR spectroscopy has demonstrated that protein structure in solution largely resembles that in a crystal and, more surprisingly, that the vast majority of copies of a protein freely floating in solution have about the same structure. Several unstructured regions (i.e. regions that have different conformations in individual molecules and/or rapidly interconvert) have recently attracted great attention (Uversky and Dunker, 2010; Vendruscolo, 2007) in protein sciences. Do we expect that most DNA/chromatin has a stable, well-defined spatial structure similar to that of proteins? How different are these structures in individual cells and how rapidly do they move around, fold and unfold?

While some specific loci may have stable conformations that are the same in all cells, we do not expect the majority of the chromatin fibers to be folded in exactly the same way in different cells. The entropic cost of ordering such gigantic molecules as chromosomal DNA in eukaryotes can run too high to achieve precise folding. Resulting significant cell-to-cell variability makes bulk methods, like chromatin capture techniques (Naumova and Dekker, 2010), discover structural characteristics averaged over millions of cells. While such variability makes it difficult to build precise 3D models, experimental data nevertheless allows to study some general features of chromatin architecture and principles that govern its organization. How can one characterize chromatin folding if no unique structure is attainable?

One productive approach offered by statistical mechanics and used in structural biology of proteins (Vendruscolo, 2007) is to consider an ensemble of conformations (not necessarily in equilibrium), found in different cells and/or at different time points during an experiment. Statistical properties of the ensemble can tell us about the principles that govern DNA packing.

Thus, the aim of building a single 3D model consistent with the measurements is replaced with the goal of finding a physical model of folding, which produces an ensemble of conformations whose properties resemble that of the ensemble studied experimentally by chromatin capture and/or optical techniques. In the search for such a model, we turn to the statistical physics of polymers that is concerned with characterizing states of a polymer that emerge as a result of interactions between the monomers, the solvent and surrounding surfaces.

Below, I will describe classical equilibrium states of the polymer and their biologically relevant and measurable statistical properties. Next, I will focus on a non-equilibrium state, the fractal globule, originally proposed in 1988 (Grosberg et al., 1988) and later suggested as a model for DNA folding inside a cell (Grosberg et al., 1993), and recently brought into the spotlight by the discovery that such a state is indeed consistent with Hi-C data obtained for human cells (Lieberman-Aiden, van Berkum et al., 2009). I will then present a summary of our recent work aimed at characterizing biophysical features the fractal globule, and the relevance of this architecture for range of biological functions. Finally, I will discuss our expectations regarding the possibility of finding the fractal globule architecture of chromatin in yeast and bacteria.

Chromatin as a polymer

The approach of statistical physics frequently deals with a coarse-grained "beads-on-a-string" representation of a polymer (Gennes, 1979; Grosberg and Khokhlov, 1994; Rubinstein and Colby, 2003). The power of this approach is that it describes an ensemble of polymer conformations that emerges at scales much greater than the size of the individual monomers and irrespective of their fine structure: whether the monomer is a single chemical group, an amino acid, or a nucleosome.

Several approximation have to be made to model chromatin fiber as a a homopolymer, i.e. a polymer with all monomers interacting in the same way, having the same size and uniform flexibility along the chain.

As a first approximation, eukaryotic chromatin can be considered as a polymer fiber formed by DNA wrapped around nucleosomes and separated by linkers of about $40-60$bp (Routh et al., 2008). This fiber has a diameter of about 10nm and a flexibility which emerges as a result of the flexibility of the linkers and partial unwrapping of nucleosomal DNA. Given that the persistence length of DNA (150bp), we obtain that about $3-4$ linkers will

provide flexibility of a persistence length of the fiber. Steric interactions between nucleosomes and possible occupancy of linkers by other DNA-binding proteins, however, can make the fiber less flexible, leading to the estimate that about $5-6$ nucleosomes form a persistence length fragment. Thus, each "bead" is not a single, but a few neighboring nucleosomes. The arrangement of neighboring nucleosomes within such a bead determines its size but is of less concern for large-scale architecture: it can be some sort of regular zig-zag pattern or an irregular blob whose fold is determined by linker lengths and nucleosome phasing (Routh et al., 2008). If the fiber is modeled as a freely-jointed chain, each segment of the chain shall have a length twice the persistence length, i.e. $10-12$ nucleosomes which corresponds to $2-2.5$Kbp of DNA. Thus a chromosome/region of 10Mb can be modeled as a chain of $4000-5000$ freely-jointed segments. Each "bead" then consists of $10-12$ nucleosomes and, depending on their arrangement, will have a volume exceeding that of its comprising DNA and histones by a factor of $3-4$, i.e. $v \approx 15 \cdot 10^3 \text{nm}^3$, allowing it to be modeled as a sphere of about $30nm$ in diameter. Alternatively, one can model chromatin as a homopolymer of the 30nm-fiber that has been observed *in vitro* but whose presence *in vivo* is debated.

The nature of interactions between the monomers remains to be discovered. These can include DNA bridging and packing by specific structural proteins, like cohesin (Nasmyth and Haering, 2009), CTCF (Phillips and Corces, 2009) or RNA molecules (Ng et al., 2007), and long-range interactions between enhancers and promoters mediated by assembly of transcription machinery (Alberts, 2008b). Interactions between the chromatin fiber and the rest of the nucleus may involve steric confinement by the lamina, anchoring to the nuclear matrix or protein-mediated bridging to the nuclear lamina (Kind and van Steensel, 2010). The power of the statistical approach is that it allows characterization of configurational ensembles that emerge as a result of these interactions while disregarding their specific features.

More complicated statistical modeling can also consider specific interactions, replacing a homopolymer with a heteropolymer of several types of beads interacting differently with each other and the lamina, e.g. regions of open and closed chromatin. One can also take into account how the local density of nucleosomes influences local flexibility of the chain and size of the "beads" (Alberts, 2008a); for example, loss of nucleosomes in a regulatory region can make it much more flexible. While it may be tempting to model bacterial chromatin as naked DNA subject to interactions mediated by DNA-bridging and structural proteins like H-NS (Fang and Rimsky, 2008) and MukBEF (Petrushenko et al., 2010), local supercoiling can lead

4

to formation of non-trivial DNA packing. The fluctuating filament model recently introduced by Wiggins, et al. (Wiggins et al., 2010) models bacterial DNA as packed into a uniform-density filament of some yet-unknown structure that is likely to include a stack of plectonemic supercoiled loops. In principle, a polymer decorated by supercoiled loops can be modeled as a branched polymer. Chromatin containing lots of crosslinks (e.g. mitotic chromosomes (Marko, 2008)) can be considered as a polymer gel.

In summary, to the first approximation, chromatin can be modeled as a homopolymer formed by DNA wrapped into nucleosomes. Such a polymer is assumed to have a constant diameter, DNA density and flexibility along the chain, with monomers experiencing excluded volume and other interactions as well as spatial confinement. More detailed models of a general branched heteropolymer may include heterogeneity of density, interactions, and shapes of the monomers.

Equilibrium states of a single polymer

The properties of a homopolymer under different conditions are presented in detail in several excellent books which should satisfy both an expert (Gennes, 1979; Grosberg and Khokhlov, 1994; Rubinstein and Colby, 2003) and a novice (Grosberg and Khokhlov, 1997) in the field. Here, we provide a quick summary, focusing on biologically relevant quantities that are be measured by optical and chromosome capture experiments. Two characteristics of primary interest are (i) the mean spatial distance $R(s)$ between two loci that are genomic distance s apart along the chain, a quantity that is measured by FISH; and (ii) the probability of contact $P_c(s)$ between two loci that are distance s apart, which can be calculated from the chromatin capture data. Both quantities are averaged over the conformational ensemble in polymer physics, and over a population of cells in an experiment.

The random coil

A polymer in which monomers that are far apart along the chain do not interact, even when approaching each other in space, is called an ideal chain. Under certain conditions, the behavior of real chains can be well-approximated by an ideal chain. Irrespective of the local mechanisms of chain flexibility (e.g. a worm-like chain, a freely jointed chain, etc.), the behavior of sufficiently long fragments of the chain resembles a 3D random walk. The characteristic size of the polymer, which can be either its root-mean-squared end-to-end distance R or its mean radius of gyration R_g, scale

with the polymer length N as

$$R(N) \sim N^{1/2}. \tag{1}$$

The end-to-end distance of a sub-chain of length s has the same scaling, i.e $R(s) \sim s^{1/2}$. Here and below, polymer length is measured in terms of the polymer's persistence length ℓ_p, i.e. $N = L/\ell_p$, which depends on the local mechanism of flexibility and for naked DNA was measured to be $\ell_p \approx 150$bp. Alternatively, one can use the Kuhn length b which is defined as a length of a bond in a freely jointed chain that has the same end-to-end distance. For the worm-like chain model, the Kuhn length is twice the persistence length $b = 2\ell_p$ and

$$R(N) = bN^{1/2}, \tag{2}$$

where $N = L/b$.

Characteristic for the ideal chain is the power $\nu = 1/2$ of $R(s) \sim s^\nu$. This scaling of the end-to-end distance with s can be tested by FISH experiments, where two loci a distance s apart are labeled and visualized in individual cells, allowing the measurement of spatial distance between them. A recent review (Emanuel et al., 2009) suggests that significant cell-to-cell variability, however, makes it hard to obtain reliable estimates of ν.

Chromosome capture methods (Lieberman-Aiden et al., 2009; Miele and Dekker, 2009; Ohlsson and Göndör, 2007; van Berkum and Dekker, 2009), in turn, can provide data on the probability of contact between loci distance s apart along the genome. For the ideal chain one can obtain

$$P_c(s) \sim s^{-3/2}. \tag{3}$$

Note that a polymer in this state is rather expanded and has a low density. For example, a random coil of the *E. coli* genome has a size of $R \approx b\sqrt{N} = b\sqrt{L/b} = \sqrt{4.6 \cdot 10^6 \cdot 300}$bp $\approx 12\mu$m, which is much greater than the size of the *E. coli* bacterium. If excluded volume interactions between the monomers are taken into account, then the scaling of the polymer size changes to $R \sim N^{3/5}$, a case referred to as the "swollen coil". The swollen coil has a size even larger than that of the random coil and is unlikely to be a relevant model for DNA packing.

The equilibrium globule

If attraction between the monomers dominates over excluded volume repulsion, or if the polymer is confined to a sufficiently small volume, the polymer

undergoes a coil-globule transition into an equilibrium globule. The size of the equilibrium globule scales with the polymer length as

$$R \sim N^{1/3}. \tag{4}$$

Hence, the volume occupied by the polymer scales linearly with polymer length: $V \sim R^3 \sim N$, i.e. monomers fill a fixed fraction of the volume, and the density of monomers $\rho \equiv N/V \sim const$ is independent of the polymer length and is uniform inside the globule. This uniform density contrasts with that of an ideal chain where the volume populated by the polymer $V \sim R^3 \sim N^{3/2}$ and the density decrease with N: $\rho \sim N/V \sim N^{-1/2}$, resulting in monomers occupying a tiny fraction of the volume of the coil.

Relevant for FISH experiments is the scaling of the end-to-end distance of a sub-chain with its length s. This scaling in the globule differs from that of the whole chain. In fact, according to the Flory theorem (Grosberg and Khokhlov, 1994), interactions of a chain in a dense melt are screened by other chains, making that chain behave like an almost ideal chain. In other words, a chain inside a globule behaves like a random walk, until the "walker" hits the boundary of the confining volume (or the boundary of the globule), and after such a "collision" the walker starts a new random walk uncorrelated with the previous one. The end-to-end distance of a sub-chain then scales as

$$R(s) \sim \begin{cases} s^{1/2} & \text{for} \ \ s \leq N^{2/3} \\ const & \text{for} \ \ s > N^{2/3}. \end{cases} \tag{5}$$

Note that a similar ideal regime of chains is observed in other dense polymer systems such as melts of many individual polymers. The contact probability for a sub-chain of an equilibrium globule scales approximately as

$$P_c(s) \sim \begin{cases} s^{-3/2} & \text{for} \ \ s \leq N^{2/3} \\ const & \text{for} \ \ s > N^{2/3}. \end{cases} \tag{6}$$

Figure 1 shows these scaling behaviors for simulated equilibrium globules. Interestingly, FISH data for yeast chromosomes labeled at the centromere and telomere show a similar roll-over into a plateau, a characteristic feature of the equilibrium globule (see Fig in (Emanuel et al., 2009; Therizols et al., 2010)).

Another important property of the equilibrium globule is its entanglement. Computer simulations (Lieberman-Aiden et al., 2009; Virnau et al., 2005) and theoretical calculations (Grosberg, 2000; Metzler et al., 2002) have demonstrated that a long polymer folded into an equilibrium globule is highly knotted. Such knots can hamper folding and unfolding processes

(Bölinger et al., 2010), making knotted conformations rare among naturally occurring protein structures (Lua and Grosberg, 2006; Virnau et al., 2006). Because of the high degree of entanglement of the globule, folding into such knotted conformation requires a polymer to thread its ends through different loops many times. Since its slithering motion is rather slow and diffusive (polymer ends move equally forward and backward), formation of the entangled equilibrium globule is a very slow process (with equilibration time $\sim N^3$).

The fractal (crumpled) globule

According to de Gennes, polymer collapse proceeds by the formation of crumples of increasing sizes: first small crumples are folded, leading to formation of an effectively thicker polymer-of-crumples, which next forms large crumples itself, etc. Grosberg, Nechaev and Shakhnovich (1988) demonstrated that this process leads to formation of a long-lived state that they called a *crumpled globule* and which was recently referred to as a *fractal globule* (here we adopt the latter notation). According to (Grosberg et al., 1988), the fractal globule is characterized by a hierarchy of crumples, thus the conformation resembles a fractal structure with a fractal dimension of 3 (for comparison, the Gaussian coil formed by an ideal chain has a fractal dimension of 2). In other words, every sub-chain (of length s) [2] is folded into a conformation akin to the equilibrium globule i.e. $R(s) \sim s^{1/3}$ (see Fig 3). These globules emerge due to topological constraints: every sufficiently long chain experiences these constraints imposed by other parts of the polymer and collapses into a globule subject to these confining interactions.

Since the fractal globule is dense, its size scales linearly with polymer length the same way the equilibrium globule does:

$$R(N) \sim N^{1/3}. \tag{7}$$

Since the fractal globule consists of globules formed on all scales (FIg 3), the scaling of the size of a sub-chain follows the same law:

$$R(s) \sim s^{1/3} \tag{8}$$

for $s > N_*$, where $N_* \sim 10$ (see [2]). Comparison of these equations with the scaling for the equilibrium globule (5) reveals two major differences: (i) the

---

[2]the sub-chain should be sufficiently long $s > N_*$, short sub-chains behave like an ideal chain(Grosberg et al., 1988). The value of N_* depends on the density (Grosberg et al., 1988), but simulations show that it can be as small as $\approx 10 - 30\ell$

scaling of the end-to-end distance has a power of 1/3 for the fractal globule, rather than 1/2 for the equilibrium globule; and (ii) the plot of $R(s)$ vs s for the fractal globule does not have a plateau present in the equilibrium globule (see Figure 1A). Such differences can be detected by high-resolution FISH experiments with averaging over a sufficiently large number of cells (e.g. (Yokota et al., 1995)), but suffecient cell-to-cell variability can make it hard to distinguish 1/2 and 1/3 (Emanuel et al., 2009).

The contact probability for the fractal globule was not computed in the original (Grosberg et al., 1988) contribution and was difficult to compute analytically without making drastic simplifications. Our group used simulations to obtain the scaling of $P_c(s)$ in the fractal globule. We use traditional Monte Carlo simulations of a polymer freely-jointed chain modeled as spherical impenetrable beads with diameter b (Imakaev and Mirny, 2010; Lieberman-Aiden et al., 2009). The simulations took care of not violating topological constrains as tested by computing Alexander polynomials on reduced chains (Virnau et al., 2006). Simulating collapse from a coil state by applying a confining spherical cage, we obtained fractal globules for chains as long as $N = 128,000$ monomers (Imakaev and Mirny, 2010). Excluded volume and topological constrains are the only factors that drive the system, allowing to avoid ambiguities caused by the choice of a specific potential of interactions and temperature. Obtained fractal globules show a robust scaling

$$P_c(s) \sim s^{-1} \tag{9}$$

for $s > N_*$. Comparison with the contact probability for corresponding equilibrium globule (6) shows a significant difference in the exponent (-1 vs $-3/2$) and the lack of a plateau for large s which is present in the equilibrium globule (Fig 1B). These features have been used in the recent analysis of the human Hi-C data (Lieberman-Aiden, van Berkum et al., 2009).

The fractal globule in human chromatin architecture

Recently, chromosomal contacts in human cells have been characterized by the Hi-C experiments (Lieberman-Aiden, van Berkum et al., 2009). Among several important observations brought to light by this study is the dependence of the contact probability $P_c^{\mathrm{exp}}(s)$ on genomic distance s:

$$P_c^{\mathrm{exp}}(s) \sim s^{\alpha}, \quad \alpha \approx -1, \tag{10}$$

9

for s in the range from 0.25Mb to about 10Mb. While the original paper made this statement based on a linear fitting of $\log P_c^{\mathrm{exp}}(s)$ vs $\log s$, our more recent analysis using other statistical techniques (e.g. the Kolmogorov-Smirnov test) demonstrated that α is indeed very close to -1 (Fundenberg and Mirny, 2010) in this and likely a broader range of distances.

The scaling of s^{-1} is easy to intuit. First, it means that loci two-fold farther apart are two-fold less likely to interact. Second, if contacts are interpreted as chromatin loops, then there is no mean or characteristic loop length: loops of all length are present and the mean is not well-defined for s^{-1} scaling. This observation contrasts with the earlier loop models of chromatin packing (Münkel et al., 1999; Sachs et al., 1995).

Another important feature of $P_c^{\mathrm{exp}}(s)$ obtained by Hi-C experiments is the lack of a clear plateau at large s, that would be indicative of the equilibrium globule (see Fig1B). While some rise in the slope is observed for $s \gtrsim 10 - 100$Mb, it is not statistically significant due to the fact that some chromosomes are shorter and the lack of dynamic range which would allow extremely low-frequency interactions to be resolved.

As explained above, fractal globule is the only ensemble of states that is consistent with (a) scaling -1 of contact probability, and (b) the lack of plateau in $P_c(s)$. The Hi-C results allows us to reject the equilibrium globule as well as the swollen or ideal coils, and any regular arrangement of open loops as a high-order structure of the chromatin in human cells. However, the fractal globule itself is not necessarily formed by 46 unconstrained chromosomes. Individual chromosomes can have free or anchored loops/domains which individually fold into fractal globules.

While FISH data largely do not have sufficient precision to discriminate between $1/2$ and $1/3$ exponents of $R(s)$ (Emanuel et al., 2009), data for large distances $s > 10Mb$ in human chromosome 4 (Yokota et al., 1995) are best fit by $R(s) \sim s^{0.32}$ (Münkel et al., 1999), which is consistent with the fractal globule's $1/3$ at this scale.

Folding, unfolding and loop opening

Beyond being the only model that fits Hi-C data, the fractal globule has several important properties that make it an attractive way of organizing chromatin in a cell.

The fractal globule is easy to form: as we showed by simulations (Imakaev and Mirny, 2010; Lieberman-Aiden et al., 2009), a non-specific collapse of a polymer naturally leads to a fractal globule conformation, provided that

10

topological constraints are in place, i.e. the chain cannot cross itself. For a chromatin fiber, such a collapse could be induced by DNA-binding condensing/linking proteins like cohesin (Nasmyth and Haering, 2009), CTCF (Ohlsson et al., 2010; Phillips and Corces, 2009) or structural RNAs (Ng et al., 2007) and can span large chromosomal domains.

The fractal globule is unentangled, i.e. it contains no knots since it maintains the topology of an open state. Dynamics of chromatin opening from the unentangled fractal conformation are very different from that of the knotted conformation of the equilibrium globule, as we demonstrated by simulations. For example, a region of about 1Mb in a fractal globule of 8Mb can easily unfold if molecular crosslinks that keep it condensed are removed (Imakaev and Mirny, 2010; Lieberman-Aiden et al., 2009). A similar region of the equilibrium globule does not fully open up as it remains trapped by multiple entanglements.

Such ability to rapidly unfold can be of great importance for gene activation which has been shown to cause de-condensation of a large $(0.5 - 2\text{Mb})$ genomic region (Müller et al., 2001). Our model suggests that such displacement/modification of crosslinking proteins/RNAs in a spatially small area of condenses loop is sufficient to trigger its large-scale decondensation. Modification or displacement of crosslinking proteins can be accomplished by some members of transcription machinery or polymerase complex that are recruited to the activated locus. This can explain why decondensation depends on the presence of transcription factor activation domains (Carpenter et al., 2005) or polymerase activity (Müller et al., 2001). In simulations, the unfolded loop can rapidly move around allowing it to sample space, finding a transcription factory (Cope et al., 2010). Since the unfolded loop has a scaling of $R(s) \sim s^{1/2}$ as compared to $R(s) \sim s^{1/3}$ in the folded state, in the unfolded state it can exceed in size a much longer domain folded into a compact fractal globule. This argument is consistent with experiments of Müller et al. (2001) which demonstrated that a $0.5\mu m$ spot decondenses into a $1 - 10\mu m$ loop.

Thus, the fractal globule architecture can provide a mechanism that allows rapid and large-scale opening of genomic loci and their spatial motion in the unfolded state. Importantly, all these events happen spontaneously in response to local removal/modification of crosslinking proteins.

The fractal globule, topological constraints and chromosomal territories

Our recent simulations (Imakaev and Mirny, 2010; Lieberman-Aiden et al., 2009) demonstrated that when a chain is folded into a fractal globule, each sequential region of the chain occupies a distinct spatial region (see Fig 2 and Fig 4). The original paper of Grosberg, et al. (Grosberg et al., 1988) suggested such spatial segregation emerges due to topological constraints on scales greater than a certain length N_*. This segregation of sub-chains is akin to the segregation of polymer rings that occurs due to topological constraints and was suggested as a mechanism that leads to formation of chromosomal territories (de Nooijer et al., 2009; Dorier and Stasiak, 2009; Rosa and Everaers, 2008; Vettorel et al., 2009). In contrast to chromosomal territories that separate chromosomes into spatially distinct regions, spatial segregation in the fractal globule occurs on all scales (Fig 3). This suggests presence of *genomic territories* where a continuous genomic region is spatially compact (Fig 3), and different regions occupy different locations (Fig 4). The fractal globule suggests the presence of genomic territories in a broad range of scales: from tens of kilobases to tens of megabases. While sub-chromosomal domains have been visualized as non-overlapping spatial entities (Visser and Aten, 1999), more systematic study of genomic territories can test predictions made by the fractal globule model.

While scaling of the contact probability suggested the presence of fractal globules for genomic regions of up to $5 - 10$ Mb long, it is possible that full chromosomes and their relative packing follow the same principle of fractal globule architecture. Several recent studies have suggested that topological constraints can lead to the emergence of chromosomal territories (de Nooijer et al., 2009; Dorier and Stasiak, 2009; Rosa and Everaers, 2008; Vettorel et al., 2009). By simulating chromosomes as polymer chains or rings of various length that are either confined to a small volume (de Nooijer et al., 2009; Dorier and Stasiak, 2009) or equilibrated in a melt of other chromosomes (Vettorel et al., 2009), these studies have observed spatial segregation of chains. Such segregation closely resembles chromosomal segregation observed by optical microscopy (Cremer and Cremer, 2010). Moreover Vettorel, Grosberg and Kremer (2009) demonstrated that polymer rings equilibrated in a high-density melt have statistical properties resembling that of the fractal globule. Dorier and Stasiak (2009) have shown that topological constraints are more important than excluded volume in inducing spatial segregation of rings. However, unrealistically short rings upon

extreme confinement used in Dorier and Stasiak (2009) necessitate further studied of this phenomenon. Rosa and Everaers (2008) examined the equilibrium and kinetics of polymer rings and chains. They report observing robust $R(s) \sim s^{1/3}$ scaling for equilibrated rings and as a transient, long-lived intermediate of confined polymer chains. They note that this scaling is consistent with $R(s) \sim s^{0.32}$ obtained for human chromosome 4 using FISH techniques (Yokota et al., 1995). Rosa and Everaers (2008) also noted that simple polymer collapse and equilibration do not lead to such fractal scaling (i.e. same scaling for all length-scales), but lead instead to the equilibrium globule.

In summary, these studies demonstrated that topological constraints, the same ones that lead to formation of the fractal globule, lead to spatial segregation of chromosomes. Segregation in the fractal globule, leads to emergence of "genomic territories" on all scales above some N_*. Biologically, this means that any region of the genome folded into a fractal globule is spatially compact, rather than spatially spread. Decondensation and spreading can be caused by either active displacement of crosslinking proteins/RNAs (e.g. during gene activation) or by violation of the topological contains (e.g. by topo II or unrepaired double-stranded DNA breaks, see below).

Mixing and crosstalk

Despite the territorial organization created by topological constraints, there is a great deal of interaction between individual regions of the fractal globule. Figure 5 presents two neighboring crumples inside a fractal globule, showing a great deal of interdigitation of the two crumples.In fact, our study demonstrated (Lieberman-Aiden, van Berkum et al., 2009) that the number of interactions $M(s)$ in a region of length s of a fractal globule scales as

$$M(s) \sim s \sim R^3(s) \sim V(s), \tag{11}$$

i.e. linearly with its *volume*, rather than its surface area. As we showed analytically, this scaling follows directly from $P(s) \sim s^{-1}$ scaling of the contact probability. This means that individual regions deeply penetrate into each other's volumes (see 5), rather than touch each other on the surface, as spheres, polyhedra or other squishy but impenetrable objects would do. In other words, a crumple of a fractal globule has a fixed (independent of its size) fraction of its volume that is involved in interactions.

Moreover, in the fractal globule the number of contacts between two crumples of lengths s_1 and s_2 ($s_{1,2} \ll N$) that are separated by a distance l

along the chain scales as

$$M_{1,2}(l) \sim \frac{s_1 s_2}{l} \sim \frac{V_1 V_2}{l}. \tag{12}$$

Thus, the number of interactions is proportional to the product of the crumples' volumes. Such penetration means a great deal of possible cross-talk between individual regions of all sizes (loci, chromosomal arms, etc.) despite their spatial segregations. Thus, the fractal globule simultaneously provides two seemingly contradictory features: spatial segregation of genomic regions on all scales and their extensive cross-talk.

Stability of the fractal globule

While providing a number of advantages, the fractal globule is a long-lived intermediate on the way to becoming an equilibrium globule. What are the factors that determine its metastability? How can cells maintain the fractal globule organization of chromatin for a long time?

The original theory of the fractal globule (Grosberg et al., 1988) suggested that (i) the lifetime of the fractal globule is determined by a time ($\sim N^3$) required to thread the ends of the polymer through the whole globule, allowing the formation of knots; (ii) a chain with attached ends (e.g. a loop or a ring) remains in the fractal globule state forever. We tested these statements by simulations demonstrating that equilibration of the fractal globule is indeed a very slow process (see 6) with the time exceeding $\sim N^3$. Rosa and Everaers estimate that it would take more than 500 years for a chromosomal fiber to equilibrate (Rosa and Everaers, 2008). Strikingly, we also found that a chain with confined ends nevertheless slowly converts into an equilibrium globule, while remaining unentangled (see Fig 6) (Imakaev and Mirny, 2010). This process is also very slow for long loops, but very short loops may be subject to such equilibration.

The lifetime of the fractal globule naturally depends on the stringency of the topological constraints. Such constraints can be violated in the cell by DNA topoisomerase II enzyme (topo II). Topo II cuts both strands of one DNA double helix, passes another unbroken DNA helix through it, and then religates the cut DNA. In doing so, it can knot and unknot DNA (Vologodskii, 2009). To test the role of topo II we made simulation where occasional strand passing were allowed. This simulation show rapid equilibration of fractal globule, suggesting that active topo II could destroy fractal globule architecture and chromosomal territories requiring some additional mechanisms for their stabilization.

It was proposed that formation and maintenance of chromosomal territories requires unviolated topological constraints (Dorier and Stasiak, 2009) leading the authors to suggest that topo II was unable to act on nucleosomed chromatin fibers. This conjecture is, however, not supported by experimental findings which demonstrated that at least *in vitro* topo II is able to act on nucleosomed DNA as efficiently as on naked DNA, reducing its positive supercoiling (Salceda et al., 2006). However, the ability of topo II to facilitate the passage of two nucleosomed chromatin fibers through each other *in vivo* as well as topo II activity and essentiality during the interphase are unknown.

Stabilization of the fractal globule can involve anchoring as well as reversible and irreversible crosslinking of DNA by proteins or RNA molecules. Simulations show that while reversible crosslinking cannot prevent eventual equilibration, it can significantly slow it down (Imakaev and Mirny, 2010).

To manifest in the cell, fractal globule and topological territories should not necessarily be stable indefinitely. They should persist at least for the duration of single cell cycle, as chromosomal architecture is re-esstablished upon mitosis. Recent photo-activation experiments beautifully demonstrated that chromosomal architecture is maintained for $10 - 15$ hours and is completely reset upon mitosis (Strickfaden et al., 2010). Mechanism that suppress stand passing and otherwise stabilizes the fractal globule as well as the rest of chromatin architecture during the interphase are yet to be established.

Other organisms

Relevance of the fractal globule architecture to chromosome organization in other organisms depends primarily on the chromosome length and available nuclear volume. The DNA density in a diploid human cell is approximately 6000Mb packed in the nuclear volume of $\approx 300 \mu m^3$, i.e. $20 Mb/\mu m^3$. Baker's yeast, in contrast has a density of about $12 Mb/3\mu m^3 = 4 Mb/\mu m^3$. Such a vast difference in density may entail very different chromatin architectures. While a significant fraction of the yeast nucleus is occupied by nucleoli that are inaccessible to the chromatin (Therizols et al., 2010), the remaining volume may be almost sufficient for a loose coil formed by an ideal or swollen coil. For example, a long yeast chromosomal arm of $N = 0.5$Mb corresponds to about $200b$ (with Kuhn length $b = 30$nm) and has a characteristic size $R(N) \approx b\sqrt{N}$ or $\approx bN^{3/5} \approx 0.4 - 0.7\mu m$ which easily fits inside the available volume of yeast nucleus and matches high-resolution FISH measurements

(Therizols et al., 2010). Chromosome capture (4C, 5C, Hi-C,noble-yeast) measurements in yeast (Duan et al., 2010) will provide critical information about the scaling of the contact probability $P(s)$ revealing whether chromatin in yeast is packed into a fractal globule or not. Topological constraints in yeast, may nevertheless lead to segregation of chromosomes into less pronounced chromosomal territories (Berger et al., 2008; Haber and Leung, 1996; Therizols et al., 2010).

Recent experiments have shed light on the organization of the bacterial chromosome. Wiggins, et al. (Wiggins et al., 2010) have demonstrated linear organization of the bacterial chromosome in *E.coli*. The origin of replication was found to be positioned close to cell center, while the two "arms" extend symmetrically. Moreover, this study demonstrated that the spatial distance between any locus and the origin goes precisely linearly with the genomic distance between the two. This organization and fluctuations in loci positions are explained by a mechanical Fluctuating Spring model which assumes regular packing of DNA inside the nucleoid. Statistical models of DNA packing in bacteria (Jun and Wright, 2010) focused on a potential mechanism of DNA segregation upon cell division, suggesting that chain entropy is sufficient for spontaneous segregation of two DNA chromosomes (Jun and Wright, 2010). A statistical polymer model that can explain the observed linear scaling of spatial and genomic distance has yet to be developed. We conjecture that a fractal globule confined to the elongated geometry of the *E.coli* nucleoid can exhibit such linear scaling due to segregation of sub-chains. Again, chromosome capture can provide data complementary to optical measurements, yielding a clearer understanding of the principles that govern folding of bacterial the chromosome.

The fractal globule, topological constraints and cancer

There are a few interesting connections between the concept of the fractal globule and cancer. From the historic work of Boveri (Boveri, 1914) to recent characterization of cancer genomes (International Cancer Genome Consortium, 2010), it has been known that cancer cells carry numerous genomic rearrangements. Chromatin structure could play a role in molecular mechanisms involved in formation of genomic rearrangements and influence the distribution of rearrangements observed in cancer.

Recent characterization of somatic copy-number alteration across many human cancers (Beroukhim et al., 2010) have provided a high-resolution map

of such events and revealed two classes of rearrangements: global, such as deletions or amplifications of a complete chromosomal arm; and focal which happen on much smaller scales. The abundance of such events and significant sample-to-sample differences in the patterns of observed alterations suggest that the vast majority of these events are passenger mutations, i.e. random genetic events. Strikingly, the frequency of an alteration (insertion or deletion) of a genomic region of length s scales as

$$f(s) \approx s^{-1} \tag{13}$$

for the range of $0.1 \leq s \leq 50$. This resembles the scaling of the probability of contact between two loci distance s apart obtained by Hi-C for human chromosomes (Lieberman-Aiden, van Berkum et al., 2009). We conjecture that these two scalings are connected: if two loci are more likely to interact, they are more likely to be subject to a recombination/repair event, i.e. deletion of the formed loop or its amplification. This way, the $1/s$ scaling in the contact probability leads to the same scaling in the frequency of genomic alteration. This hypothesis is now being further tested in our lab.

Another interesting connection between the fractal globule and cancer stems from the fact that double-stranded DNA breaks can lead to strand passing and hence to violation of topological constraints. Double-stranded breaks are widespread in certain forms of cancer and are produced by deficiencies of repair and recombination machineries (Weinberg, 2007). Topological constraints, on the other hand, are central for the maintenance of the fractal globule and chromosomal territories. Abundant double-stranded breaks are likely to cause partial opening of domains folded into fractal globules leading to some degree of chromosome decondensation. Note that if the equilibrium globule or a random coil were the state of the chromatin, double-stranded breaks would have no effect. Consistent with these conjectures are experimental findings of local chromatin decondensation at the sites of double-stranded breaks (Kruhlak et al., 2006) and global chromatin decondensation upon malignant transformation (Ye et al., 2001). Such decondensation, in turn can help cancer to reverse chromatin condensation and gene silencing associated with cell differentiation (Weinberg, 2007). Thus, cancer cells may use double-stranded breaks as one of the mechanisms of de-differentation.

Double-stranded breaks can also lead to faster equilibration of the globule and melting of the boundaries of chromosomal territories. Note that the Hi-C data(Lieberman-Aiden, van Berkum et al., 2009) discussed above were obtained for two cancer cell lines (GM06990 and K562), both showing contact probabilities characteristic for the fractal globule. Further chromosome

capture and fluorescence microscopy experiments on cells subject to different levels of double-stranded induction treatment can test these predictions of the fractal globule model.

Summary and outlook

Introduced about 20 years ago (Grosberg et al., 1988) and proposed then as a model for DNA packing in the cell (Grosberg et al., 1993; Grosberg et al., 1988), the concept of the fractal globule is an attractive model of chromatin organization during interphase in human cells. It is the only statistical polymer model that is consistent with chromosome capture data, providing experimentally observed $P(s) \sim s^{-1}$ scaling (Lieberman-Aiden, van Berkum et al., 2009) ; and the only model that fits the FISH scaling of $R(s) \sim s^{0.32}$ (Rosa and Everaers, 2008; Yokota et al., 1995). The span of genomic lengths over which the fractal globule persists has yet to be established, as chromosome capture data fit the fractal globule for $0.1 \lesssim s \lesssim 10$Mb, while FISH data has close to 1/3 scaling on longer scales $s \gtrsim 10$Mb (Rosa and Everaers, 2008). High-resolution single-molecule single-cell microscopy methods may be able to overcome current limitations of the FISH method caused, in part, by significant cell-to-cell variability of spatial distances.

Several biophysical properties of the fractal globule make it a particularly appealing model of chromatin organization.

- The fractal globule is formed spontaneously due to topological constraints by chromatin condensation and is able to maintain its topological state for a long time.

- By virtue of being largely unknotted, any region of the fractal globule can easily and rapidly unfold and translocate, becoming accessible to transcription and other protein machinery of the cell.

- Folding into the fractal globule leads to formation of genomic territories (Fig 2), i.e. a conformation where any specific genomic locus is folded into compact crumples (Fig 3), and distinct loci occupy distinct spatial locations in the nucleus (Fig 4). Despite this territorial organization, folded loci form a very large number of interactions with each other (Fig 5, with the number of interactions proportional to the product of the volumes of interacting crumples). When expanded to the scale of whole chromosomes, these features of the fractal globule correspond to chromosomal territories and suggest extensive crosstalk between the chromosomes.

18

- The fractal globule is a long-lived intermediate that gradually decays into an equilibrium globule (Fig 6) that lacks many of the properties of the fractal globule and is not consistent with the experimental data. Activity of the topo II enzyme significantly accelerates this process, while crosslinking by proteins like CTCF slows it down.

Many of these properties, predicted theoretically and observed in simulations, can be tested experimentally to better characterize the state of the chromatin inside a cell. For example, genomic territorial organization can be tested using high-resolution optical microscopy by methods like PALM or STORM (Betzig et al., 2006; Rust et al., 2006).

The role of topo II in the organization of the interface of chromosomes is intriguing. Its ability to facilitate passage of nucleosomed chromosomal fibers, thus violating topological constraints, can be tested experimentally. Similarly, stability of chromosomal/genomic territories to activity of topo II in vivo can be assayed. Chromatin pulling experiments (Marko, 2008) can help to test the degree of DNA entanglement and to characterize contributions of topological constraints and crosslinking proteins to the folded state. These questions are central to understanding the role of topological constraints in the formation and support of chromosomal organization during the interphase (de Nooijer et al., 2009; Dorier and Stasiak, 2009; Lieberman-Aiden et al., 2009; Rosa and Everaers, 2008; Vettorel et al., 2009). Chromosome capture methods (Dekker, 2008) can reveal how chromatin is organized in different organisms and different tissues, as well as observe the evolution of its structural state upon differentiation or malignant transformation.

Obtaining the precise structure of chromatin akin to the structure of a folded protein may not be feasible as chromatin structures can differ significantly from cell to cell. However, approaches based on statistical physics of polymers and high-quality experimental measurements can help characterize the state of the chromatin as a conformational ensemble, revealing basic organizing principles behind chromatin folding and dynamics.

Acknowledgments

References

Alberts B (2008a) *Molecular biology of the cell*, 5th ed edn, Garland Science, New York.

Alberts B (2008b) *Molecular biology of the cell: Chapter 7*, 5th ed edn, Garland Science, New York.

Berger A. B , Cabal G. G , Fabre E , Duong T , Buc H , Nehrbass U , Olivo-Marin J.-C , Gadal O and Zimmer C (2008) High-resolution statistical mapping reveals gene territories in live yeast, *Nat Methods* **5**: 1031–7.

Beroukhim R , Mermel C. H , Porter D , Wei G , Raychaudhuri S , Donovan J , Barretina J , Boehm J. S , Dobson J , Urashima M , Mc Henry K. T , Pinchback R. M , Ligon A. H , Cho Y.-J , Haery L , Greulich H , Reich M , Winckler W , Lawrence M. S , Weir B. A , Tanaka K. E , Chiang D. Y , Bass A. J , Loo A , Hoffman C , Prensner J , Liefeld T , Gao Q , Yecies D , Signoretti S , Maher E , Kaye F. J , Sasaki H , Tepper J. E , Fletcher J. A , Tabernero J , Baselga J , Tsao M.-S , Demichelis F , Rubin M. A , Janne P. A , Daly M. J , Nucera C , Levine R. L , Ebert B. L , Gabriel S , Rustgi A. K , Antonescu C. R , Ladanyi M , Letai A , Garraway L. A , Loda M , Beer D. G , True L. D , Okamoto A , Pomeroy S. L , Singer S , Golub T. R , Lander E. S , Getz G , Sellers W. R and Meyerson M (2010) The landscape of somatic copy-number alteration across human cancers, *Nature* **463**: 899–905.

Betzig E , Patterson G. H , Sougrat R , Lindwasser O. W , Olenych S , Bonifacino J. S , Davidson M. W , Lippincott-Schwartz J and Hess H. F (2006) Imaging intracellular fluorescent proteins at nanometer resolution, *Science* **313**: 1642–5.

Bölinger D , Sułkowska J. I , Hsu H.-P , Mirny L. A , Kardar M , Onuchic J. N and Virnau P (2010) A stevedore's protein knot, *PLoS Comput Biol* **6**: e1000731.

Boveri T (1914) *Zur Frage der Entstehung Maligner Tumoren (Concerning the origin of malignant tumours)*, The Company of Biologists.

Carpenter A. E , Memedula S , Plutz M. J and Belmont A. S (2005) Common effects of acidic activators on large-scale chromatin structure and transcription, *Mol Cell Biol* **25**: 958–68.

Cope N. F , Fraser P and Eskiw C. H (2010) The yin and yang of chromatin spatial organization, *Genome Biol* **11**: 204.

Cremer T and Cremer M (2010) Chromosome territories, *Cold Spring Harb Perspect Biol* **2**: a003889.

de Nooijer S , Wellink J , Mulder B and Bisseling T (2009) Non-specific interactions are sufficient to explain the position of heterochromatic chromocenters and nucleoli in interphase nuclei, *Nucleic Acids Res* **37**: 3558–68.

Dekker J (2008) Gene regulation in the third dimension, *Science* **319**: 1793–4.

Dorier J and Stasiak A (2009) Topological origins of chromosomal territories, *Nucleic Acids Res* **37**: 6316–22.

Duan Z , Andronescu M , Schutz K , McIlwain S , Kim Y. J , Lee C , Shendure J , Fields S , Blau C. A and Noble W. S (2010) A three-dimensional model of the yeast genome, *Nature* **465**: 363–7.

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**: 025008.

Fang F. C and Rimsky S (2008) New insights into transcriptional regulation by h-ns, *Curr Opin Microbiol* **11**: 113–20.

Fundenberg G and Mirny L. A (2010) Statistical properties of the chromosome capture data for human and yeast cells. to be published.

Gennes P. G. d (1979) *Scaling concepts in polymer physics*, Cornell University Press, Ithaca, N.Y.

Grosberg (2000) Critical exponents for random knots, *Phys Rev Lett* **85**: 3858–61.

Grosberg A , Rabin Y , Havlin S and Neer A (1993) Crumpled globule model of the three-dimensional structure of dna, *EPL (Europhysics Letters)* **23**: 373.

Grosberg A. I and Khokhlov A. R (1997) *Giant molecules: here, and there, and everywhere–*, Academic Press, San Diego.

Grosberg A. Y and Khokhlov A. R (1994) *Statistical physics of macromolecules*, AIP Press, New York.

Grosberg A. Y , Nechaev S and Shakhnovich E (1988) The role of topological constraints in the kinetics of collapse of macromolecules, *Journal de Physique* **49**: 2095–2100.

Haber J. E and Leung W. Y (1996) Lack of chromosome territoriality in yeast: promiscuous rejoining of broken chromosome ends, *Proc Natl Acad Sci U S A* **93**: 13949–54.

Imakaev M and Mirny L. A (2010) Fractal globule as model of chromatin architectute: biophysical properties to be published.

International Cancer Genome Consortium (2010) International network of cancer genome projects, *Nature* **464**: 993–8.

Jun S and Wright A (2010) Entropy as the driver of chromosome segregation, *Nat Rev Microbiol* **8**: 600–7.

Kind J and van Steensel B (2010) Genome-nuclear lamina interactions and gene regulation, *Curr Opin Cell Biol* .

Kruhlak M. J , Celeste A , Dellaire G , Fernandez-Capetillo O , Müller W. G , McNally J. G , Bazett-Jones D. P and Nussenzweig A (2006) Changes in chromatin structure and mobility in living cells at sites of dna double-strand breaks, *J Cell Biol* **172**: 823–34.

Lieberman-Aiden E , van Berkum N. L , Williams L , Imakaev M , Ragoczy T , Telling A , Amit I , Lajoie B. R , Sabo P. J , Dorschner M. O , Sandstrom R , Bernstein B , Bender M. A , Groudine M , Gnirke A , Stamatoyannopoulos J , Mirny L. A , Lander E. S and Dekker J (2009) Comprehensive mapping of long-range interactions reveals folding principles of the human genome, *Science* **326**: 289–93.

Lua R. C and Grosberg A. Y (2006) Statistics of knots, geometry of conformations, and evolution of proteins, *PLoS Comput Biol* **2**: e45.

Marko J. F (2008) Micromechanical studies of mitotic chromosomes, *Chromosome Res* **16**: 469–97.

Metzler R , Hanke A , Dommersnes P. G , Kantor Y and Kardar M (2002) Equilibrium shapes of flat knots, *Phys Rev Lett* **88**: 188101.

Miele A and Dekker J (2009) Mapping cis- and trans- chromatin interaction networks using chromosome conformation capture (3c), *Methods Mol Biol* **464**: 105–21.

Müller W. G , Rieder D , Kreth G , Cremer C , Trajanoski Z and McNally J. G (2004) Generic features of tertiary chromatin structure as detected in natural chromosomes, *Mol Cell Biol* **24**: 9359–70.

Müller W. G , Walker D , Hager G. L and McNally J. G (2001) Large-scale chromatin decondensation and recondensation regulated by transcription from a natural promoter, *J Cell Biol* **154**: 33–48.

Münkel C , Eils R , Dietzel S , Zink D , Mehring C , Wedemann G , Cremer T and Langowski J (1999) Compartmentalization of interphase chromosomes observed in simulation and experiment, *J Mol Biol* **285**: 1053–65.

Nasmyth K and Haering C. H (2009) Cohesin: its roles and mechanisms, *Annu Rev Genet* **43**: 525–58.

Naumova N and Dekker J (2010) Integrating one-dimensional and three-dimensional maps of genomes, *J Cell Sci* **123**: 1979–88.

Ng K , Pullirsch D , Leeb M and Wutz A (2007) Xist and the order of silencing, *EMBO Rep* **8**: 34–9.

Ohlsson R , Bartkuhn M and Renkawitz R (2010) Ctcf shapes chromatin by multiple mechanisms: the impact of 20 years of ctcf research on understanding the workings of chromatin, *Chromosoma* .

Ohlsson R and Göndör A (2007) The 4c technique: the 'rosetta stone' for genome biology in 3d?, *Curr Opin Cell Biol* **19**: 321–5.

Petrushenko Z. M , Cui Y , She W and Rybenkov V. V (2010) Mechanics of dna bridging by bacterial condensin mukbef in vitro and in singulo, *EMBO J* **29**: 1126–35.

Phillips J. E and Corces V. G (2009) Ctcf: master weaver of the genome, *Cell* **137**: 1194–211.

Rosa A and Everaers R (2008) Structure and dynamics of interphase chromosomes, *PLoS Comput Biol* **4**: e1000153.

Routh A , Sandin S and Rhodes D (2008) Nucleosome repeat length and linker histone stoichiometry determine chromatin fiber structure, *Proc Natl Acad Sci U S A* **105**: 8872–7.

Rubinstein M and Colby R. H (2003) *Polymer physics*, Oxford University Press, Oxford.

Rust M. J , Bates M and Zhuang X (2006) Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (storm), *Nat Methods* **3**: 793–5.

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 U S A* **92**: 2710–4.

Salceda J , Fernández X and Roca J (2006) Topoisomerase ii, not topoisomerase i, is the proficient relaxase of nucleosomal dna, *EMBO J* **25**: 2575–83.

Strickfaden H , Zunhammer A , van Koningsbruggen S , Köhler D and Cremer T (2010) 4d chromatin dynamics in cycling cells, *Nucleus* **1**: 1–14.

Therizols P , Duong T , Dujon B , Zimmer C and Fabre E (2010) Chromosome arm length and nuclear constraints determine the dynamic relationship of yeast subtelomeres, *Proc Natl Acad Sci U S A* **107**: 2025–30.

Uversky V. N and Dunker A. K (2010) Understanding protein non-folding, *Biochim Biophys Acta* **1804**: 1231–64.

van Berkum N. L and Dekker J (2009) Determining spatial chromatin organization of large genomic regions using 5c technology, *Methods Mol Biol* **567**: 189–213.

Vendruscolo M (2007) Determination of conformationally heterogeneous states of proteins, *Curr Opin Struct Biol* **17**: 15–20.

Vettorel T , Grosberg A. Y and Kremer K (2009) Statistics of polymer rings in the melt: a numerical simulation study, *Phys Biol* **6**: 025013.

Virnau P , Kantor Y and Kardar M (2005) Knots in globule and coil phases of a model polyethylene, *J Am Chem Soc* **127**: 15102–6.

Virnau P , Mirny L. A and Kardar M (2006) Intricate knots in proteins: Function and evolution, *PLoS Comput Biol* **2**: e122.

Visser A. E and Aten J. A (1999) Chromosomes as well as chromosomal subdomains constitute distinct units in interphase nuclei, *J Cell Sci* **112 (Pt 19)**: 3353–60.

Vologodskii A (2009) Theoretical models of dna topology simplification by type iia dna topoisomerases, *Nucleic Acids Res* **37**: 3125–33.

Weinberg R. A (2007) *The biology of cancer*, Garland Science, New York.

Wiggins P. A , Cheveralls K. C , Martin J. S , Lintner R and Kondev J (2010) Strong intranucleoid interactions organize the escherichia coli chromosome into a nucleoid filament, *Proc Natl Acad Sci U S A* **107**: 4991–5.

Ye Q , Hu Y. F , Zhong H , Nye A. C , Belmont A. S and Li R (2001) Brca1-induced large-scale chromatin unfolding and allele-specific effects of cancer-predisposing mutations, *J Cell Biol* **155**: 911–21.

Yokota H , van den Engh G , Hearst J. E , Sachs R. K and Trask B. J (1995) Evidence for the organization of chromatin in megabase pair-sized loops arranged along a random walk path in the human g0/g1 interphase nucleus, *J Cell Biol* **130**: 1239–49.

Figures
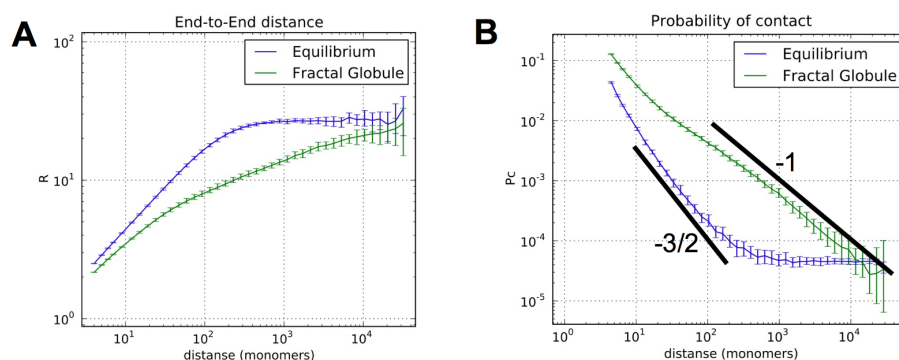
Figure 1: (A) Root-mean squared end-to-end distance $R(s)$ as a function of the genomic distance s between the ends of a subchain (in the units of ℓ) for globules of $N = 32,000$ monomers. Blue: equilibrium globule, green: fractal globule. At small s, both globules show scaling characteristic of the self-avoiding random walk $(3/5)$, followed by $1/2$ of the ideal coil. Notice there is a plateau for the equilibrium globule. (B) Probability of a contact as a function of genomic distance s for the equilibrium globule (blue) and the fractal globule (green). Noticethe robust scaling of -1 which spans two orders of magnitude for the fractal globule.
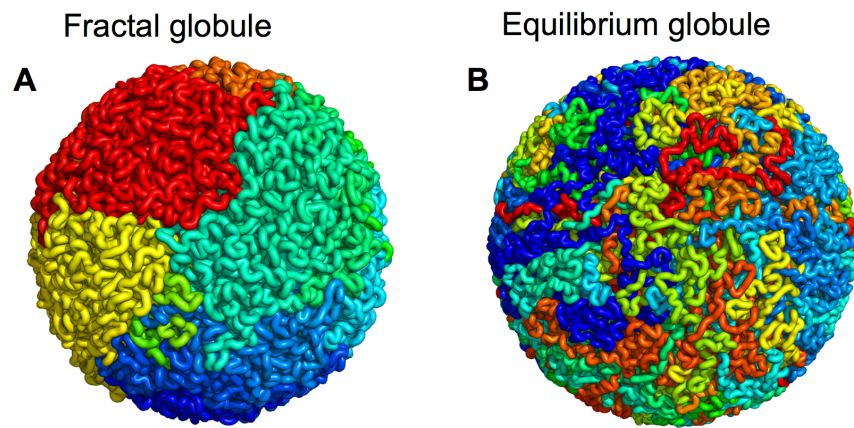
Figure 2: Conformations of the fractal (A) and equilibrium (B) globules. The chain is colored from red to blue in rainbow colors (C). The fractal globule has a striking territorial organization which strongly contrasts with the mixing observed in the equilibrium globule.
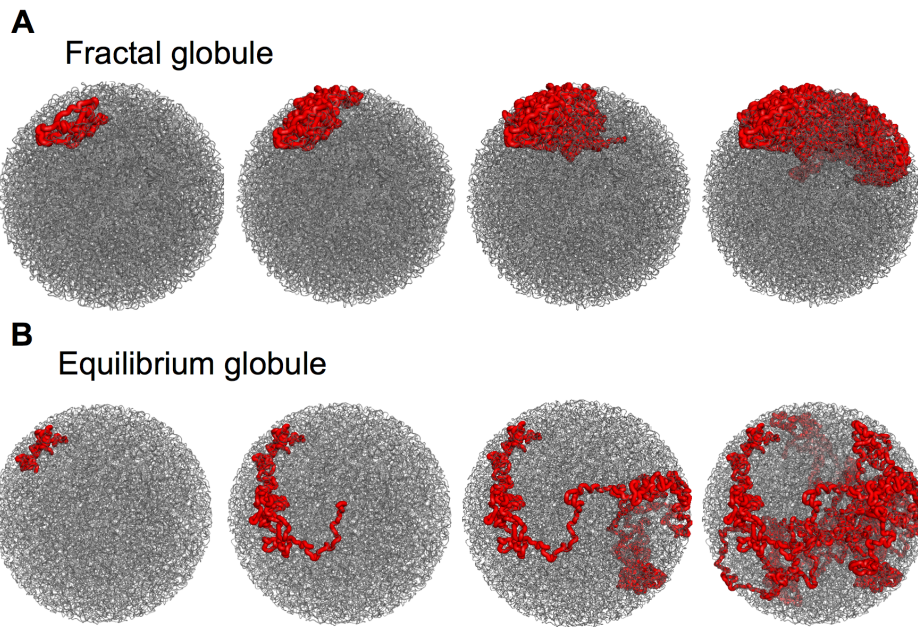
Figure 3: The fractal globule (A) consist of dense globules formed on all scales. Subchains of 100, 300, 1000 and 3000 monomers (left to right) are shown by a red tube in a globule of $N = 32,000$ monomers. For comparison, same regions of the equilibrium globule (B) are diffuse inside the globule.
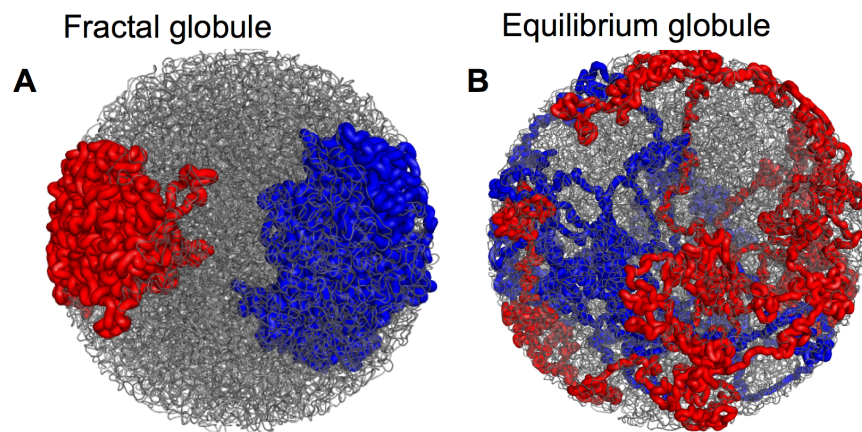
Fractal globule

A

Equilibrium globule

B

Figure 4: Territorial organization of the fractal globule (A) is evident when two chains of 1000 monomers each are outlined. The equilibrium globule,in contrast, has two chains mixed together in space.
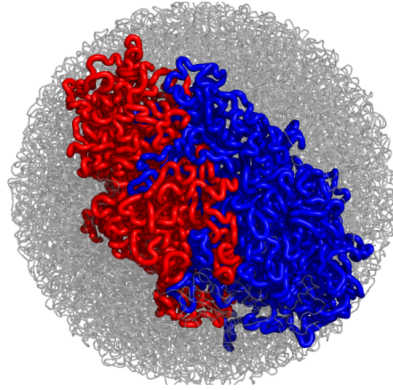
Figure 5: Despite having an organized territorial architecture, spatially neighboring regions of the fractal globule (shown in red and blue) have a large number of interactions between then, deeply penetrating into each others volumes. The number of interactions a crumples has scales linear with its volumes (see eq (12)). Thus a fixed fraction of crumples volume (rather than its surface) is involved in interactions.
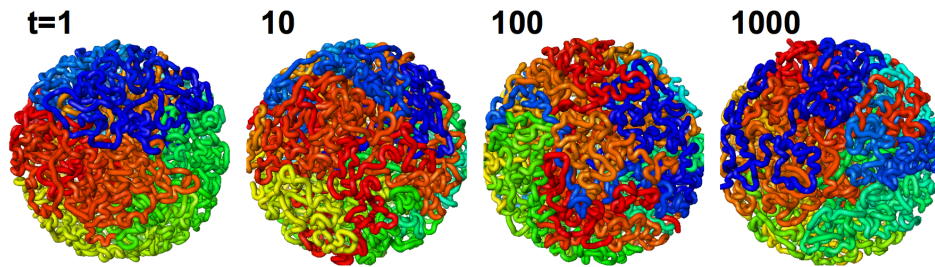
Figure 6: Equilibration of the fractal globule. A series of snapshots obtained at four logarithmically spaced timepoints of long equilibration simulations. Notice gradual loss of the territorial organization, characteristic of the fractal globule, and increasing mixing, leading to formation of the equilibrium globule. Since the ends of the globule remain attached to the surface while being able to slide on it, the structure remains unentangled. This equilibration is very slow. The details of these simulations will be published elsewhere.