Structure *14*, 197–204, February 2006 DOI 10.1016/j.str.2005.10.016

Modeling a Self-Avoiding Chromatin Loop: Relation to the Packing Problem, Action-at-a-Distance, and Nuclear Context

Michaël Bon,[1,4] Davide Marenduzzo,[2,3,4] and Peter R. Cook[1,*]
[1]Sir William Dunn School of Pathology
University of Oxford
South Parks Road
Oxford, OX1 3RE
United Kingdom
[2]Mathematics Institute University of Warwick
Coventry, CV4 7AL
United Kingdom
[3]The Rudolf Peierls Centre for Theoretical Physics
University of Oxford
Keble Road
Oxford, OX1 3NP
United Kingdom

Summary

There is now convincing evidence that genomes are organized into loops, and that looping brings distant genes together so that they can bind to local concentrations of polymerases in "factories" or "hubs." As there remains no systematic analysis of how looping affects the probability that a gene can access binding sites in such factories/hubs, we used an algorithm that we devised and Monte Carlo methods to model a DNA or chromatin loop as a semiflexible (self-avoiding) tube attached to a sphere; we examine how loop thickness, rigidity, and contour length affect where particular segments of the loop lie relative to binding sites on the sphere. Results are compared with those obtained with the traditional model of an (infinitely thin) freely jointed chain. They provide insights into the packing problem (how long genomes are packed into small nuclei), and action-at-a-distance (how firing of one origin or gene can prevent firing of an adjacent one).

Introduction

Pro- and eukaryotic genomes are positioned within cells. At the global level, whole bacterial chromosomes are often oriented so that their replication origins lie nearer the poles than the midline (Gitai et al., 2005), and gene-rich chromosomal segments in human fibroblasts are found nearer the nuclear center than gene-poor ones (Bolzer et al., 2005). Proximity to other active genes also influences whether an origin fires, or a gene is transcribed (Emerman and Temin, 1986; Lucas et al., 2000; Cook, 2003; Jun et al., 2004). Moreover, there is now convincing evidence that genomes are organized into loops, and that looping plays an important role in controlling gene activity by bringing distant genes together so that they can bind to local concentrations of relevant proteins (Cook, 2003; Chambeyron and Bickmore, 2004). For example, replicating DNA is looped by attachment to "factories" containing DNA polymerases at the membrane in bacteria and the nuclear interior in eukaryotes (Lemon and Grossman, 1998; Cook, 1999). In addition, different genes are transcribed in association with various nuclear "bodies," including rDNA and tRNA genes with nucleoli, histone genes with Cajal bodies, globin genes with "active chromatin hubs" or factories, and the major histocompatibility complex with promyelocytic leukemia bodies (de Laat and Grosveld, 2003; Spector, 2003; Osborne et al., 2004). As these various structures contain high concentrations of polymerases and associated factors, it seems likely that rates of replicational and transcriptional initiation will be governed by position in the loop and the frequency with which genes can access binding sites on the structures.

Although loops have been modeled previously (Levene and Crothers, 1986; Marko and Siggia, 1995; Merlitz et al., 1998; Rippe, 2001; Vilar and Leibler, 2003; Ballaeff et al., 2004; Mergell et al., 2004; Vilar and Saiz, 2005), there remains no systematic analysis of how gene position affects the probability of accessing a binding site in such structures. Several interrelated problems complicate such modeling. First, the structure of chromatin in vivo is unknown (Woodcock and Dimitrov, 2001). Second, a loop contains too many constituents to be modeled at the atomic level, so simplifications must be applied; for example, 20 kbp of (unlooped) nucleosomal DNA was modeled as a zig-zagging string of oblate ellipsoids, but the spacing and angles between ellipsoids profoundly affect results (Wedemann and Langowski, 2002; Mergell et al., 2004). Third, the popular approaches used to model DNA have important deficiencies in this context. Thus, those involving freely jointed or worm-like chains (De Gennes, 1979) work well when applied to naked double-stranded DNA (where polymer persistence length is much greater than thickness); however, their application to structures like a chromatin fiber (where the persistence length of 40–100 nm is roughly comparable to the thickness of the 30 nm fiber; Katritch et al., 2000) is problematic. They also have the drawback that the chains are infinitely thin, but we require that no two segments of the loop occupy the same volume in and around the binding site in the crowded cell where the DNA density is high (Luby-Phelps, 2000; Minton, 2001). Self-avoidance is usually introduced by forcing the chain to follow a nonintersecting path along a regular lattice, or by modeling nucleosomes as impenetrable beads connected by springs (Katritch et al., 2000; Wedemann and Langowski, 2002; Mergell et al., 2004).

Here, we report a new, to our knowledge, method to explicitly model a DNA or chromatin loop as a semiflexible (self-avoiding) tube (Gonzalez and Maddocks, 1999; Marenduzzo and Micheletti, 2003) attached to a sphere representing a factory/hub/body. Unlike models involving freely jointed or worm-like chains, ours is not easily tractable analytically. Therefore, we apply Monte Carlo procedures and examine how loop thickness, rigidity, and contour length affect where particular segments of the loop lie relative to binding sites on the sphere.

*Correspondence: peter.cook@path.ox.ac.uk
[4]These authors contributed equally to this work.

Results—which are compared with those obtained by modeling an infinitely thin chain (see Experimental Procedures)—provide insights into the packing problem (how long genomes are packed into small nuclei), action-at-a-distance (how activity of one gene influences that of neighboring genes), and nuclear context (why genes are active in certain positions and not others).

Results and Discussion

The Approach

As discussed in the Introduction, loops of DNA (in prokaryotes) or chromatin (in eukaryotes) are often tethered to structures called factories, hubs, or bodies. These structures contain high concentrations of polymerases and their associated factors, so it seems likely that the rates of initiation of replication and transcription will be governed by the frequency with which genes in the loop access binding sites on the structure. Therefore, we model chromatin (and DNA) as a self-avoiding tube of appropriate thickness that is tethered at its ends to a sphere (Figure 1). Starting with a circular tube, successive, randomly chosen moves are used to generate a contorted loop; after creating many equivalent structures, the probability that a particular promoter or origin occupies a specific volume is calculated. This probability reflects the way the point diffuses through space. Tube length in nm can be converted to kbp assuming 1, 5, and 8 nucleosomes (200, 1000, and 1400 bp DNA) are packed into 11 nm of the linear length of 11, 20, and 30 nm tubes, respectively. However, values in kbp should be treated cautiously, as packing ratios are unknown (Bystricky et al., 2004); the trends seen will have greater significance.

The Emptiness of Outer Space and the Packing Problem

Figure 2A shows where points at different distances, *s*, along the bottom half of a loop are likely to be relative to the center of the sphere; values for the top half are identical, as the loop is symmetric. The structure modeled is a typical chromatin loop attached to a transcription factory in a dividing HeLa cell (Cook, 1999). Zig-zagging models for chromatin have supplanted those involving 30 nm solenoids (Woodcock and Dimitrov, 2001), so tube diameter was set at 20 nm to reflect a wider zig-zagging fiber that can interpenetrate to some extent. As expected, no point on the loop is ever found within the sphere or the surrounding shell that extends a tube radius from the surface (i.e., $R + \Delta = 47.5$ nm; Figure 1B). The 34 nm point on the loop is concentrated toward the outer limit of its accessible volume (i.e., $R + \Delta + 34 = 81.5$ nm; Figure 2A); the high density near the surface limits folding, so the tube tends to start straight out from the sphere. Conversely, the 417 nm point is rarely seen beyond 200 nm (Figure 2A); once the tube extends into less crowded space, it tends to fold, and few structures exist with the needed long and straight arms. In other words, nine-tenths the volume that is accessible in theory is essentially completely empty. This packing occurs spontaneously, without extra energy cost. As we shall demonstrate below, this emptiness of outer space allows close packing of loops associated with different spheres to reduce significantly the scale of the

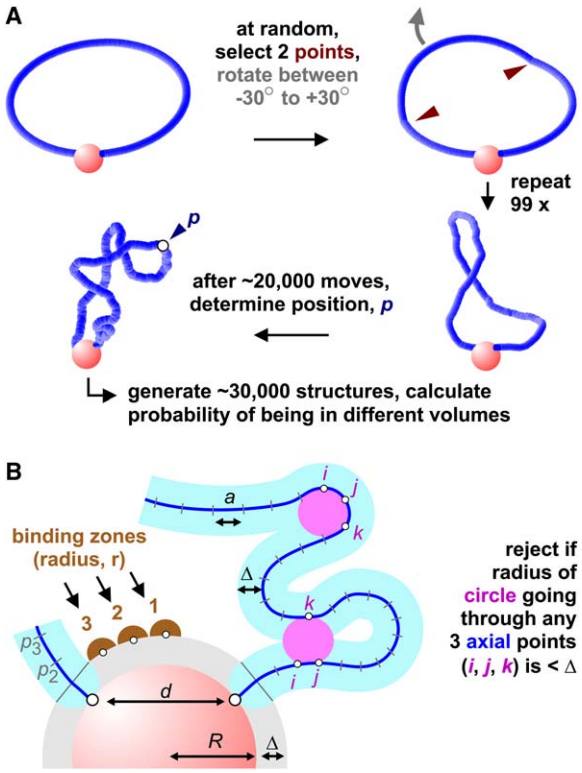

Figure 1. The Approach

(A) A circular tube of uniform thickness (capped at each end by a hemisphere) is tethered to a sphere. Two points on the tube are selected at random, and the intervening segment rotated at a randomly chosen angle between −30° and +30°. If any part of the tube finds itself in space occupied by other parts of the tube or the sphere, the move is rejected; this process is repeated until an acceptable move is found. After ~20,000 successful moves, the tube has adopted a random looped structure, and the position in 3D space of a particular point, *p*, is determined. This process is repeated to generate ~30,000 such structures, and the average position of *p* is calculated. Structures are visualized with "Rasmol" software.

(B) Rejection criteria and characteristics of tube and binding zone. Points, p_1, p_2, ...p_n are scattered along the tube axis. Two portions of the tube (length, *L*; radius, Δ) do not occupy the same space if parts of the axis are separated by $\geq 2\Delta$. This is monitored by determining if radii of circles going through any triplet of axial points *i*, *j*, *k* are $\geq \Delta$; then, tight bends incompatible with both tube thickness and too close an approach of distance portions of the tube are prevented. Tube thickness prevents the axis approaching within Δ of the surface of the sphere (gray zone). Positions of three binding zones are shown.

"packing" problem—how a long contour length of DNA might be folded into a tiny nucleus (in man, ~2 m are packed into ~10 μm).

Points between the 34 and 417 nm points have intermediate distributions, but increasing distance has progressively less effect on the volume occupied by most chromatin. Thus, the peak in probability of the 77 nm point is significantly further out than that of the 34 nm point (Figure 2A); space is still crowded close to the sphere. However, points between 162 and 417 nm have essentially the same distributions; all share the same (emptier) space (Figure 2A). Shorter and longer loops yield similar trends. Thus, the 34 nm points in loops with contour lengths of 425–2550 nm occupy essentially the same space (data not shown), which is reflected by the constant distance between the center of the sphere

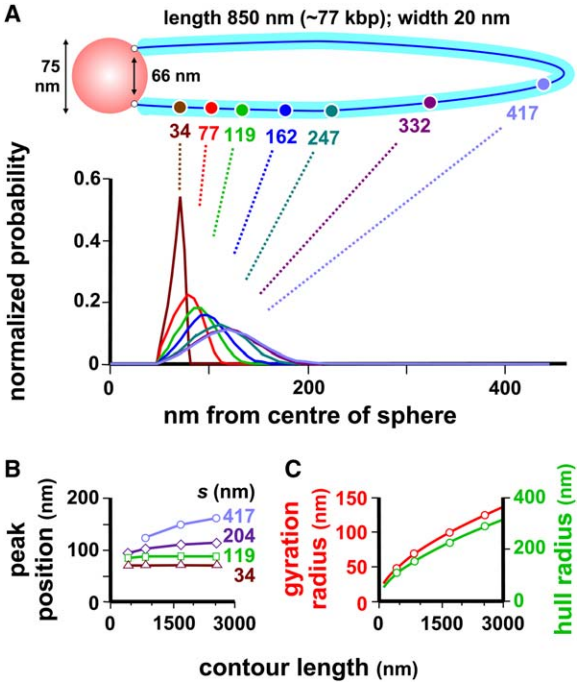

Figure 2. Loops Spontaneously Pack Close to the Sphere

(A) Emptiness of outer space. The drawing illustrates a "standard" loop attached to a sphere; the probabilities that different points on the centerline of the loop (from 34 to 417 nm) are found within successive 8.5 nm shells around the sphere are shown in the graph (bottom). For detailed parameters concerning this and other loops, see Experimental Procedures. All points on the loop pack close to the surface of the sphere; most of the accessible volume is beyond 200 nm and is essentially empty.

(B) Dependence of position of peak probability relative to the sphere center (measured with curves like those in [A] for loops of different length). Peak probabilities for 34 nm points on loops of different lengths are all found at roughly the same distance from the center of the sphere; in contrast, those for the 417 nm points increase asymptotically as contour length increases.

(C) Radius of gyration of loops with different lengths, and spherical hull radius. Both gyration and hull radius (which reflects the volume of sphere plus loop) are much smaller than half the contour length; this means that each loop occupies a tiny fraction of the accessible volume, and this allows close packing of other spheres and their loops. Data points lie on the curve $f(L) = a + bL^{0.6}$, where $a = 8.08$ and $b = 1.05$ (red line), and $a = 10.26$ and $b = 2.49$ (green line).

and the peak in probability (Figure 2B). For the 417 nm points the distance between sphere center and peak increases asymptotically as contour length increases (Figure 2B). The small radius of gyration of the loop (De Gennes, 1979) and spherical hull radius (which reflects the volume of both sphere and loop; see Experimental Procedures) are both much smaller than half the contour length, and so consistent with each loop occupying a tiny fraction of the accessible volume (Figure 2C). Both quantities vary with loop length, L, according to the law $a + bL^{0.6}$, where a and b are constants (Figure 2C, legend); this law is typical of a self-avoiding walk.

Inert, "Hot," and "Cold" Zones

The probability of being found close to the surface of the sphere is of special interest, as it may determine function. Thus, the frequency of collision of origins and promoters with polymerases concentrated in the factory/hub/body to which the loop is tethered probably influences initiation frequencies, and so overall rates of replication and transcription (Cook, 1999). The curve in the lower half of Figure 3A illustrates the probabilities that points on the loop shown above it (which is the same as that modeled in Figure 2A) are found within a binding zone close to the surface. Consider the lower half of the loop (the pattern in the top half is again identical, as the binding zone lies symmetrically between tethering points). As expected, points with tethers too short to reach the binding zone (i.e., in segment *a*, which is shown in black in the illustration in Figure 3A) are never found within it; this may underlie the phenomena of "origin exclusion" and "transcriptional interference," where firing of one origin or gene prevents replication or transcription of an adjacent one (Emerman and Temin, 1986; Lucas et al., 2000; Cook, 2003; Jun et al., 2004). It has always been difficult to explain such action-at-a-distance, but here we see it is an inevitable consequence of the organization; tethering one gene to a polymerase in a factory/hub (with consequent replication or transcription of the tethered gene) prevents an adjacent gene in segment *a* from binding to the same factory/hub (and so initiating). Moreover, our predictions are in rough accord with experimental results. For example, our segment *a* extends several kbp, which compares favorably with observations that firing of one origin prevents firing of another within 4–5 kbp (Jun et al., 2004), and that transcription

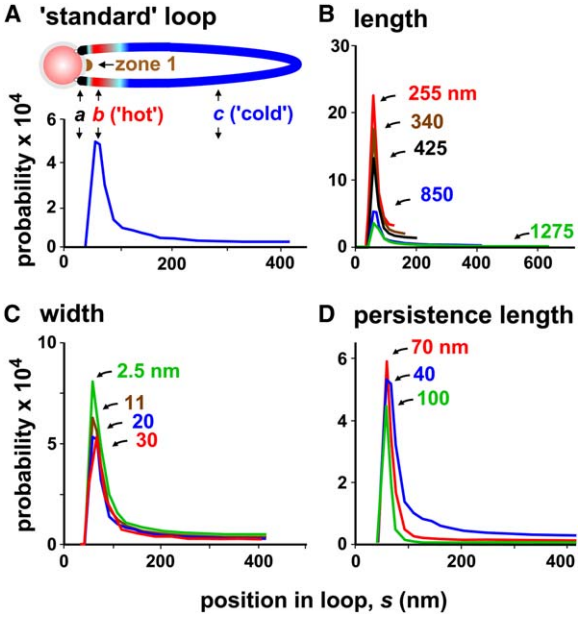

Figure 3. Binding Probabilities

(A) The illustration demonstrates the standard loop illustrated in Figure 2A, and the probabilities that different points at different fractional distances, *s*, along the loop are found within binding zone 1 are shown in the graph (bottom). Inert (*a*), hot (*b*), and cold (*c*) zones are indicated; the hot and cold zones may be eu- and heterochromatic, respectively.

(B–D) Effects of variations in tube length, width, and persistence length on the probability that different positions along the standard loop occupy binding zone 1. Peak probabilities are the lowest for long, wide, and rigid loops; therefore, heterochromatin is likely to be out in long loops as thick and rigid fibers.

of one gene prevents transcription of another lying about the same distance away (Emerman and Temin, 1986) (note, however, that the length of *a* is sensitive to modeling details, and especially to the distance between tethering point and binding zone [see below]).

In contrast to segment *a*, the next segment (i.e., *b*) can access the binding zone with high probability (Figure 3A); however, the probability soon falls off as one goes further out along the loop (i.e., into segment *c* in Figure 3A). Therefore, we might expect potentially active genes to be found in the "hot" segment *b*, and rarely active ones in "cold" segment *c* (see also the discussion of eu- and heterochromatin below). As at least 20% and 28% of the known genes on human chromosomes 21 and 22 lie within 1 and 10 kbp of another (Adachi and Lieber, 2002), position would then play a significant role in determining whether binding of one gene inhibits or stimulates activity of an adjacent one. These results show that different segments in the loop have very different (albeit low) probabilities of binding.

Similar inert, hot, and cold segments are seen as the contour length of the loop is varied (Figure 3B).

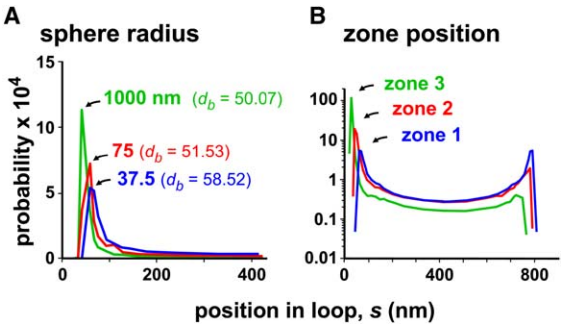

Figure 4. Effect of Sphere Radius and Position of Binding Zone

(A) The effect of sphere radius on the probability that different positions along the standard loop occupy binding zone 1 (see Figure 1B). Peak probabilities increase as radius increases; therefore, active genes are likely to be attached to large structures.

(B) Effect of placing the binding zone in the positions shown in Figure 1B on the probability that different positions along the standard loop occupy the binding zone. Positions on both arms of the loop are plotted. Proximity to the binding zone has a marked effect (compare the heights of the two peaks on each of the three curves).

Tube Thickness and Stiffness

We next varied tube thicknesses between 2.5 and 30 nm (the diameters of hydrated double-stranded DNA and inert chromatin), and focused on the effects of self-avoidance, as they cannot be analyzed with freely jointed chains. We tuned the bending rigidity of the tube so that its persistence length was the same for all thicknesses considered; we also tuned sphere radius so that $R + \Delta$—which determines the volume excluded to the tube centerline—remained constant at 47.5 nm. Then, self-avoidance alone underlies the effects seen (see Experimental Procedures). All thicknesses yield the same general pattern, but thinner tubes give peaks slightly closer to the sphere, as they can make sharper turns (Figure 3C). Moreover, probabilities fall as thickness increases (Figure 3C); they reflect the ratio of the number of configurations with a point in the binding zone relative to all possible configurations, and increasing thickness reduces the former more than the latter (as crowding near the surface increases relatively more than that in outer space). It is to be noted that the 2.5 nm curve is indistinguishable from that given by a (zero thickness) worm-like chain with persistence length of 40 nm; however, the 11 nm curve is clearly different (data not shown). Stiffening the tube by increasing persistence length also decreases probabilities in the cold region by more than an order of magnitude (Figure 3D); bending stiffer tubes back to access the binding zone costs more energy. In contrast, peaks generally become larger and narrower; stiffer tubes are straighter and have increased probabilities of accessing the binding zone, but numbers fall relatively to all possible configurations (the particular discretization length [*a*; Figure 1B] and position in the loop [*s*] chosen for analysis influence the precise value of the peak probability [see Experimental Procedures]; here, they combine to yield a value for a persistence length of 100 nm that is lower than that for 70 nm [Figure 3D]).

Eu- and Heterochromatin

We have seen that distance, thickness, and rigidity all conspire together to minimize access to the binding zone. We would then predict that the most active parts of the genome (i.e., euchromatin) would be found in segment *b*, and the least active parts (i.e, heterochromatin) out in thick and rigid fibers in segment *c*. Various results (also obtained with HeLa cells) are consistent with these predictions (reviewed by Cook, 1999): (1) detaching most of the loop with nucleases preferentially removes inactive genes; (2) heterochromatin is typically composed of dense, thick, fibers; (3) the distant cold segment constitutes ~75% of the loop, and this compares with the observed value of 68%–84% chromatin being heterochromatic (Kimura and Cook, 2001).

What might trigger heterochromatinization of distant segments? Nucleosomal concentration falls off with distance from the sphere, so infrequent contacts between nucleosomes in outer space could trigger the alteration in histone code that promotes aggregation.

Large Spheres and Nuclear Context

Increasing sphere radius increases the probability that "hot" segments access the binding zone (Figure 4A); the distance from tethering point to binding zone (d_b) decreases as the curvature of the sphere decreases, and d_b affects the probability (see below). However, there is also a competing effect: increasing sphere radius decreases the probability that cold segments are found in the binding zone (Figure 4A), and this is again due to the increased crowding near the sphere surface (this effect is captured with the freely jointed chain [Figure 5A]). As a result, there is a trade-off between increasing peak probabilities and decreasing distal ones.

These results help explain why the most active genes always seem to be associated with large structures. Consider, for example, replication in a mammalian cell, which begins in small factories that grow progressively during S phase (Cook, 1999). Growth in factory size will then progressively increase the probabilities that even poor origins in the hot zone can initiate, to ensure complete genome duplication by the end of S phase. Moreover, we might expect larger factories to possess more

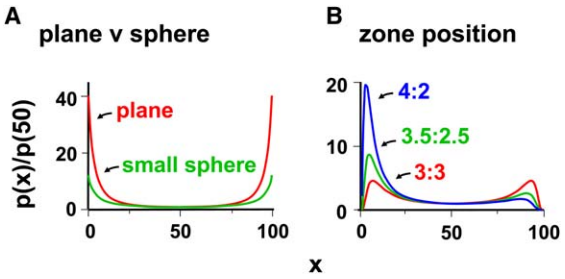

Figure 5. Binding Probabilities for Some Freely Jointed Chains

(A) Effects of varying sphere radius. The probability that a point at position x along a loop ($N = 100$) tethered to a sphere of radius negligible with respect to the radius of gyration, or to a sphere with infinite radius (i.e., plane), occupies a binding zone that is symmetrically placed between tethering points is given. Probabilities have been scaled relative to that found for a point (p[50]) at position N/2. It can be seen that the increase in volume excluded to the chain when it is attached to a plane renders the segments close to the center of the loop consistently "colder." This effect is also seen with the self-avoiding tube.

(B) Effect of varying position of binding zone. The probability—scaled as in (A)—that a point at position x along a loop ($N = 100$) tethered to a sphere of infinite radius (plane) occupies different binding zones is given. Distances between the binding zone and each tethering point are indicated (total distance is 6 units). The relative positioning of tethering and binding zones affects peak height, but less than in the case of a self-avoiding tube (compare with Figure 4B).

binding sites than do smaller ones, and this would increase peak probabilities even further. A similar effect may explain why bacterial replication factories are placed at the cell membrane (Lemon and Grossman, 1998)—an origin attached to what is effectively a sphere of infinite radius has an increased probability of binding. Moreover, the most highly transcribed eukaryotic genes are also associated with large structures rich in the needed polymerases (e.g., rDNA genes associate with nucleoli, histone genes with Cajal bodies, globin genes in erythroid cells with large transcription factories; Spector, 2003; Osborne et al., 2004); promoter occupancy of the binding zone increases as radius increases to ensure the maximum rate of transcription. However, a promoter in a cold segment is less likely to be found in the binding zone (Figure 4A), and so there may be a trade-off between increasing peak probabilities and decreasing distal ones, which may result in a compromise over factory radius.

Position of Binding Zone

Precisely which point in the hot segment has the highest probability depends on tube width, persistence length, and sphere radius (Figures 3B–3D and 4A); it is also very sensitive to distance between tethering points (not shown) and position of the binding zone. For example, moving the zone progressively closer to one tethering point progressively increases peak probabilities in the nearest arm, while reducing those in the distant arm—as might be expected (Figure 4B; positions along both arms of the loop are illustrated, and probabilities are plotted logarithmically). Strikingly, however, distant segments are colder than in the symmetric case. This effect is not captured by the freely jointed chain, and is due to steric hindrance between segments close to the tethering point (compare Figures 4B and 5B).

Conclusions

We employed a new, to our knowledge, algorithm and Monte Carlo methods to model a loop of DNA (or chromatin) attached to a sphere (representing a factory/hub/body) in two ways—as a flexible tube of finite thickness (Figure 1), and as an infinitely thin freely jointed chain (see Experimental Procedures). The tube is clearly a closer approximation of the biological structure, and it yields results that are qualitatively different from those obtained with the traditional approach (Figures 3C and 4B illustrate some of the most obvious differences). Moreover, fitting looping probabilities to freely jointed or worm-like chains may lead to misleading values for persistence length; for example, two tubes with different thickness but similar stiffness (Figure 3C) would yield different persistence lengths. In contrast, fitting the probabilities to the tube should yield information on all parameters defining the tube and sphere. Our results also provide insights into the packing problem (how long genomes are packed into small nuclei), action-at-a-distance (how activity of one gene influences that of neighboring genes), and nuclear context (why genes are active in certain positions and not others). We hope that, as detailed information on loop structure becomes available, these results will enable us to make quantitative predictions about gene activity.

Experimental Procedures

A loop is modeled both as a self-avoiding tube and as a freely jointed chain (a worm-like chain with small bending rigidity is expected to behave in the same manner).

The Self-Avoiding Tube

We model a loop as a semiflexible tube (contour length, L) of finite thickness 2Δ (Gonzalez and Maddocks, 1999; Marenduzzo and Micheletti, 2003) and bending rigidity K_b. Finite thickness impacts on two distinct features of polymer conformation. First, it constrains the local radius of curvature to be not less than Δ to avoid singularities. Second, it prevents any distant portions of the tube from occupying the same volume; the tube is self-avoiding, and centerlines of any two portions must be separated by more than 2Δ (Figure 1B). Self-avoidance is achieved by ensuring that radii of circles going through any triplet of points on the discretized centerline is greater than Δ (Figure 1B; Gonzalez and Maddocks, 1999; Maritan et al., 2000; Marenduzzo and Micheletti, 2003). Other implementations of the thickness constraint include the approximation of the tube as a succession of suitably joined cylinders (Rawdon, 2000); this produces smoother curves at the same level of discretization, but requires more computational time, and so has not been adopted here.

One is thus led to consider the Hamiltonian for the unconstrained self-avoiding tube:

$$H(\Gamma) = \sum_{ijk} V_3(r_{ijk}), \tag{1}$$

where V_3 is the three-body potential used to enforce tube thickness (Gonzalez and Maddocks, 1999; Maritan et al., 2000; Marenduzzo and Micheletti, 2003) and Γ is the configuration of the tube. The argument of V_3 is the radius of the circle going through the triplet of distinct points i, j, k, and V_3 has the form

$$V_3(r) = \begin{cases} 0 & \text{if } r > \Delta, \\ \infty & \text{otherwise} \end{cases}. \tag{2}$$

In order to model a loop attached to an impenetrable sphere, tube ends were modeled as hemispheres whose centers were constrained to lie on the surface of a sphere of radius $R + \Delta$, and points on the tube centerline were not allowed to lie within $R + \Delta$ (Figure 1B).

In all cases, the potential we consider includes a term for bending rigidity which reads

$$H_b(\Gamma) = -\frac{K_b}{a}\sum_i \cos(\theta_i), \quad (3)$$

where θ_i is the angle between the ith and (i + 1)th discretization segment on the centerline, and a is the length of one such segment. A typical discretization of an 850 nm tube involved 100 points on its centerline.

Persistence Length

Δ and K_b determine persistence length, ξ, which obeys the following approximate equation (Marenduzzo and Micheletti, 2003):

$$a\xi^{-1} = -\log\left(1 - \frac{a}{\beta K_b} + \frac{a^2}{2\Delta^2}\frac{1}{\exp\left(\frac{\beta K_b a}{2\Delta^2}\right) - 1}\right), \quad (4)$$

with $\beta = 1/T$ (the Boltzmann constant is taken to be 1).

Two noteworthy limiting behaviors are obtained when either the bending rigidity term or the thickness constraint is absent. When the thickness constraint is absent and $K_b >> a$ (continuum limit), one obtains the traditional result: $\xi = \beta K_b$. The other limit corresponds to the persistence length of a discretized thick tube with no bending rigidity, and yields:

$$a\xi^{-1} = -\log\left(1 - \frac{a^2}{4\Delta^2}\right). \quad (5)$$

Therefore, even in the absence of an explicit bending penalty, the restrictions imposed by Δ on the local radius of curvature are sufficient to produce an effective persistence length.

For completeness, we add torsional rigidity, K_t, to the model,

$$\xi = \frac{9\Delta^2}{16\beta K_t}\left[1 + \sqrt{1 - \left(\frac{8\beta K_t}{3\Delta}\right)^2}\right], \quad (6)$$

which is valid for $a \to 0$ and $K_t < K_t^*$ (where $K_t^* = 3\Delta/8\beta$). If $K_t > K_t^*$, an oscillating phase is found in which the tube is locally arranged in a helix.

Monte Carlo Simulation

The classical Metropolis-Monte Carlo procedure is used to generate randomly a statistically relevant set of representative configurations at temperature T (Sokal, 1996). Starting with an arbitrary configuration (in our case, a circular tube attached to a sphere), a new one is created randomly, and the difference (ΔE) between the two energies computed. The new configuration is accepted with a probability depending on ΔE, according to the Metropolis test (Sokal, 1996). This procedure is then repeated (typically around 50 million Monte Carlo moves were performed to calculate the binding probability for a single position on the loop).

The new configuration is generated with the crankshaft move (Sokal, 1996). A segment of the loop is selected at random, and rotated through a randomly chosen angle (typically from $-\pi/6$ to $\pi/6$). This is the only move that allows the loop constraint to be preserved during Monte Carlo dynamics. If only local crankshaft moves are chosen, the tube remains unknotted; if global moves (i.e., involving rotation of the chain between distant points) are allowed, it is known that knots can form. We verified that looping probabilities were the same, within the error, in both cases.

In order to simulate the rare occupation of a binding zone by a point on the tube, we slightly modified the algorithm of Podtelezhnikov and Vologodskii (2000) to fit our case. We considered several concentric spheres centered around the binding site with increasing radius, R_i, with i = 1, …, M (M is the number of spheres used in the calculation). The largest sphere radius, R_M, is chosen large enough that the point is always in the sphere during the whole simulation. Alternatively, we could have used a parallel tempering algorithm with a fictitious potential attracting the point to the binding site (data not shown).

After data collection, we evaluated the correlation length and calculated the statistical error accordingly. We calculated the binding probability of several points along the loop, each characterized by its fractional distance, s. Typically, about 30 positions were considered for each loop. In order to get the accuracy shown in Figure 3 (within 5%), simulations took around 80 hr of CPU time on an α computer.

Parameters Used during Modeling

Parameters used for Figure 2 were as follows: for Figure 2A, dimensions (nm) were L = 850, 2Δ = 20, ξ = 40, bending constant = 0, a = 8.5, R = 37.5; for Figure 2B, dimensions (nm) were L = 425, 850, 1700, and 2550 (other values as in A); for Figure 2C, the radius of gyration, R_g, of a polymer consisting of N beads sitting at positions $\{\vec{r}_i\}_{i=1,\ldots,N}$, and whose center of mass is in $\vec{r}_{cm}$, is calculated via the formula:

$$R_g = \left\langle \frac{\sum_{i=1}^{N}(\vec{r}_i - \vec{r}_{cm})^2}{N} \right\rangle^{1/2}, \quad (7)$$

where the brackets denote the ensemble average over all chain conformations. The hull radius considers the system formed by the sphere plus the chain, and is defined as the radius of the minimal sphere that encloses all the chain plus sphere and is centered at the sphere origin.

Parameters used for Figures 3 and 4 were generally as in Figure 2A. Binding zones (r = 8.5 nm) are placed (see Figure 1B) at positions 1 (Figures 3A–3D and 4A) or 1–3 (Figure 4B). Zone 1 is placed symmetrically between tethering points with $b_1 = b_2$ = 51.36 nm (where b_1 is distance along the surface of a gray sphere with radius R + Δ from p_1 on the left arm to the center of the zone, and b_2 is the corresponding value from p_1 on the right arm). For zone 2, b_1 = 38.29 nm and b_2 = 64.45 nm. For zone 3, b_1 = 24.65 nm and b_2 = 78.08 nm. In Figure 3B, ratios of maximum:minimum probabilities found were 6.97, 9.07, 9.42, 18.62, and 29.37 for L = 255, 340, 425, 850, and 1275 nm, respectively. Both maximum and minimum probabilities decrease as the contour length increases. Moreover, moving the tethering points closer together decreases the difference in peak height (see Figure 6 for the case of an infinitesimally thin chain; in particular, compare Figures 6A and 6B). In Figure 3C, K_b/a was set at 0, 0, 5.485, and 5.55 for tubes with diameters of 30, 20, 11, and 2.5 nm, respectively. Maximum:minimum probabilities found were 20.87, 18.62, 17.23, and 15.98 for tubes with diameters of 30, 20, 11, and 2.5 nm, respectively. In Figure 3D, the value of the persistence length lies within 5% of that indicated. K_b/a was set at 0, 7, and 11 for ξ = 40, 70, and 100 nm, respectively. Maximum:minimum probabilities found were 18.62, 48.23, and 88.47 for ξ = 40, 70, and 100 nm, respectively. Note that the 100 nm curve has a bump just beyond the peak, which cannot be seen clearly in Figure 3D at the scale used. This bump is also seen with a worm-like chain at 3–4 times the persistence length. For Figure 4A, b_1 (which equals b_2) was set at 51.36, 43.85, and 41.94 nm for R = 37.5, 75, and 1000 nm, respectively. Maximum:minimum probabilities found were 18.62, 44.46, and 182.36 for R = 37.5, 75, and 1000 nm, respectively. For Figure 4B, maximum:minimum probabilities found for the left- and right-hand peaks were: for zone 1, 18.62 and 18.62; for zone 2, 72.66 and 7.31; for zone 3, 714.68 and 2.47.

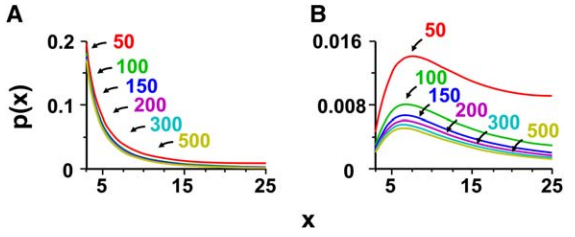

Figure 6. Nonscaled Binding Probabilities for a Symmetric Loop of Freely Jointed Chains with Different Contour Lengths Tethered to a Sphere of Infinite Radius

(A) Case (1): $\vec{r} = \vec{r}' = \vec{b}$ The curves are almost independent of contour length close to the tethering points (N = 50–500).

(B) Case (2): $\vec{r} \neq \vec{r}' \neq \vec{b}$ (see text for exact values). The colder regions in the loop (i.e., away from the peak) show a strong dependence on contour length (N = 50–500).

The Freely Jointed Chain

A loop attached to a sphere is modeled here as a freely jointed chain; the loop is infinitesimally thin and adopts a random walk. This approximation renders the model analytically tractable, and enables us to isolate the effects of self-avoidance that are included in our model. We consider the special case of a semiflexible fiber (also known as a worm-like chain) in which no bending rigidity is associated with fiber conformation (see, for example, Dekker et al., 2002). If sphere radius is much smaller than the polymer (fiber) gyration radius, R_g, the probability of having an infinitesimally thin random walk joining two attachment points in three dimensional space, which we call, respectively, $\vec{r} = (x, y, z)$ and $\vec{r}' = (x', y', z')$, is given by (Hughes, 1995):

$$P(\vec{r}, r'; N) = \frac{\exp\left(\frac{-3|\vec{r}-\vec{r}'|^2}{2Nb^2}\right)}{\left(\sqrt{2\pi Nb^2/3}\right)^3}, \tag{8}$$

where b is the Kuhn (statistical segment) length and N is the number of Kuhn lengths making up the contour length of the fiber. We also introduce a binding site at position $\vec{b} = (x_b, y_b, z_b)$.

Changing Sphere Radius

If the sphere radius, R, is much smaller than the radius of gyration ($R_g \propto N|^{0.6}b$), we find the binding probability (the probability that the point of coordinate $s = nb$, with $1 < n < N$, along the loop attaching $\vec{r}$ to $\vec{r}'$, sits at the binding site $\vec{b}$) is:

$$\begin{aligned} P(n; \vec{b}; N, R \to 0) &= \frac{1}{\left(\sqrt{2\pi nb^2/3}\right)^3} \frac{1}{\left(\sqrt{2\pi (N-n)b^2/3}\right)^3} \\ &\propto \frac{1}{n^{3/2}(N-n)^{3/2}}. \end{aligned} \tag{9}$$

This simplification occurs because the sphere is very small, so that all points must be close together and $\vec{r} \approx \vec{r}' \approx \vec{b}$. Note the probability of a point in the loop occupying the binding zone (the looping probability) has the general shape:

$$P \propto \frac{1}{n^c (N-n)^c}, \tag{10}$$

where c is 1.5 in a three-dimensional random walk.

Equation (9) is only valid if sphere radius is small in comparison with the gyration radius. For a finite sphere, one should solve the diffusion equation (random walk) with impenetrable boundaries on a sphere. For our purpose, it is easier to consider the case of a sphere of infinite radius (i.e., a plane), which, without loss of generality, we take to be the plane $z = 0$. Physically, this calculation should well approximate the case where the gyration radius is now much smaller than the sphere radius. In this case, the random walk cannot go below the plane $z = 0$ and the binding probability is (following Hughes, 1995, we take $z = z' = z_b = b$):

$$\begin{aligned} P(n; \vec{b}; N, R \to \infty) \propto \frac{(1-\exp(-6/n))}{(n)^{3/2}} \frac{(1-\exp(-6/(N-n)))}{(N-n)^{3/2}} \\ \times \exp\left(\frac{-3|\vec{r}-\vec{b}|^2}{2nb^2}\right) \exp\left(\frac{-3|\vec{r}'-\vec{b}|^2}{2(N-n)b^2}\right). \end{aligned} \tag{11}$$

In order to compare with Equation (9), it is useful to set $\vec{r} = \vec{r}' = \vec{b}$, in which case we obtain:

$$P(n; \vec{b}; N, R \to \infty) \propto \frac{(1-\exp(-6/n))}{(n)^{3/2}} \frac{(1-\exp(-6/(N-n)))}{(N-n)^{3/2}}. \tag{12}$$

For very large n and $N - n$ (close to the center of the loop), this binding probability is given by Equation (10), with c = 2.5. In the case of a plane, this means that small loops are more favored relative to long loops (compared with the case of the small sphere, in which c = 1.5).

In Figure 5A, we compare the probability, p(x), that a point x on the loop occupies a binding zone—which is symmetrically placed between the tethering points (for $\vec{r} = \vec{r}' = \vec{b}$)—for a plane and a very small sphere; trends are similar to those found with the self-avoiding tube (Figure 4B) (values for the binding probabilities are normalized relative to that found for the center of the loop [p(50)]).

Changing position of binding zone

Figure 5B illustrates how the binding probabilities for $\vec{r} \neq \vec{r}' \neq \vec{b}$ and a plane (see Equation [11]) change with location of the binding zone. Specifically we took $\vec{r} = (0, 0, b)$, $\vec{r}' = (6b, 0, b)$, and varied $\vec{b}$ along the line joining $\vec{r}$ and $\vec{r}'$. The binding probability at the exact center of the loop is unaffected by the position of the binding zone, while the peaks are strongly affected. The peak relative to the tethering point closer to the binding zone is sensibly enhanced with respect to the other, even for a mild asymmetry in binding zone position. This is to be contrasted to the behavior found for the self-avoiding loop (see main text and Figure 4B), where the peak relative to the tethering point closest to the binding zone does become larger as the binding zone becomes closer (as expected); however, the magnitude of this effect is enhanced with the self-avoiding loop. This trend is reversed away from the peak, and the binding probability becomes smaller, for asymmetrically positioned binding zones and self-avoiding loops. We ascribe this surprising effect to the fact that, away from the peak, the fiber needs to loop between tethering and binding zone, and self-avoidance of the fiber hinders binding, disallowing a large number of "bound" conformations. It thus appears that displacing the binding zone can be a way of re-adjusting the absolute binding probabilities of different points along the loop in a nontrivial way.

Changing Contour Length Of The Loop

Figure 6 illustrates how the binding probability that a point at position x on the chain (for the case of a plane) changes with the contour length of the loop, where N is varied from 50 to 500). Tethering points are as in Figure 5, and binding zones at $\vec{r} = \vec{r}' = \vec{b}$ and $\vec{b} = (3b, 0, b)$ are shown in Figures 5A and 5B, respectively. It can be seen that the scaling near the peaks deteriorates if the distance between the two tethering points increases.

Acknowledgments

We thank the Engineering and Physical Sciences Research Council for support, the Oxford Supercomputing Centre for computer resources, and the "ChromatinSim" group for running our programs.

Received: May 12, 2005
Revised: September 19, 2005
Accepted: October 4, 2005
Published: February 10, 2006

References

Adachi, N., and Lieber, M.R. (2002). Bidirectional gene organization: a common architectural feature of the human genome. Cell *109*, 807–809.

Ballaeff, A., Loudella, C.R., Mahdevan, L., and Schulten, K. (2004). Modelling DNA loops using continuum and statistical mechanics. Philos. Trans. R. Soc. Lond. A *362*, 1355–1371.

Bolzer, A., Kreth, G., Solovei, I., Koehler, D., Saracoglu, K., Fauth, C., Muller, S., Eils, R., Cremer, C., Speicher, M.R., and Cremer, T. (2005). Three-dimensional maps of all chromosomes in human male fibroblast nuclei and prometaphase rosettes. PLoS Biol. *3*, e157.

Bystricky, K., Heun, P., Gehlen, L., Langowski, J., and Gasser, S.M. (2004). Long-range compaction and flexibility of interphase chromatin in budding yeast analyzed by high-resolution imaging techniques. Proc. Natl. Acad. Sci. USA *101*, 16495–16500.

Chambeyron, S., and Bickmore, W.A. (2004). Does looping and clustering in the nucleus regulate gene expression? Curr. Opin. Cell Biol. *16*, 256–262.

Cook, P.R. (1999). The organization of replication and transcription. Science *284*, 1790–1795.

Cook, P.R. (2003). Nongenic transcription, gene regulation and action at a distance. J. Cell Sci. *116*, 4483–4491.

De Gennes, P.G. (1979). Scaling Laws in Polymer Physics (Ithaca, NY: Cornell University Press).

de Laat, W., and Grosveld, F. (2003). Spatial organization of gene expression: the active chromatin hub. Chromosome Res. *11*, 447–459.

Dekker, J., Rippe, K., Dekker, M., and Kleckner, N. (2002). Capturing chromosome conformation. Science *295*, 1306–1311.

Emerman, M., and Temin, H.M. (1986). Quantitative analysis of gene suppression in integrated virus vectors. Mol. Cell. Biol. *6*, 792–800.

Gitai, Z., Thanbichler, M., and Shapiro, L. (2005). The choreographed dynamics of bacterial chromosomes. Trends Microbiol. *13*, 221–228.

Gonzalez, O., and Maddocks, J.H. (1999). Global curvature, thickness, and the ideal shapes of knots. Proc. Natl. Acad. Sci. U.S.A. *96*, 4769–4773.

Hughes, B.D. (1995). Random Walks and Random Environments, Vol. 1: Random Walks (New York, NY: Oxford University Press).

Jun, S., Herrick, J., Bensimon, A., and Bechhoefer, J. (2004). Persistence length of chromatin determines origin spacing in *Xenopus* early-embryo DNA replication: quantitative comparisons between theory and experiment. Cell Cycle *3*, 223–229.

Katritch, V., Bustamante, C., and Olson, V.K. (2000). Pulling chromatin fibers: computer simulations of direct physical micromanipulations. J. Mol. Biol. *295*, 29–40.

Kimura, H., and Cook, P.R. (2001). Kinetics of core histones in living human cells: little exchange of H3 and H4 and some rapid exchange of H2B. J. Cell Biol. *153*, 1341–1353.

Lemon, K.P., and Grossman, A.D. (1998). Localization of bacterial DNA polymerase: evidence for a factory model of replication. Science *282*, 1516–1519.

Levene, S.D., and Crothers, D.M. (1986). Ring closure probabilities for DNA fragments by Monte Carlo simulation. J. Mol. Biol. *189*, 61–72.

Lucas, I., Chevrier-Miller, M., Sogo, J.M., and Hyrien, O. (2000). Mechanisms ensuring rapid and complete DNA replication despite random initiation in *Xenopus* early embryos. J. Mol. Biol. *296*, 769–786.

Luby-Phelps, K. (2000). Cytoarchitecture and physical properties of cytoplasm: volume, viscosity, diffusion, intracellular surface area. Int. Rev. Cytol. *192*, 189–221.

Marenduzzo, D., and Micheletti, C. (2003). Thermodynamics of DNA packaging inside a viral capsid: the role of DNA intrinsic thickness. J. Mol. Biol. *330*, 485–492.

Maritan, A., Micheletti, C., Trovato, A., and Banavar, J.R. (2000). Optimal shapes of compact strings. Nature *406*, 287–290.

Marko, J.F., and Siggia, E.D. (1995). Statistical mechanics of supercoiled DNA. Phys. Rev. E Stat. Phys. Plasmas Fluids Relat. Interdiscip. Topics. *52*, 2912–2938.

Mergell, B., Everaers, R., and Schiessel, H. (2004). Nucleosome interactions in chromatin: fiber stiffening and hairpin formation. Phys. Rev. E Stat. Nonlin. Soft Matter Phys. *70* (*1 Pt 1*), 011915, Epub July 29, 2004.

Merlitz, H., Rippe, K., Klenin, K.V., and Langowski, J. (1998). Looping dynamics of linear DNA molecules and the effect of DNA curvature: a study by Brownian dynamics simulation. Biophys. J. *74*, 773–779.

Minton, A.P. (2001). The influence of macromolecular crowding and macromolecular confinement on biochemical reactions in physiological media. J. Biol. Chem. *256*, 10577–10580.

Osborne, C.S., Chakalova, L., Brown, K.E., Carter, D., Horton, A., Debrand, E,, Goyenechea, B., Mitchell, J.A., Lopes, S., Reik, W., and Fraser, P. (2004). Active genes dynamically co-localize to shared sites of ongoing transcription. Nat. Genet. *36*, 1065–1071.

Podtelezhnikov, A.A., and Vologodskii, A.V. (2000). Dynamics of small loops in DNA molecules. Macromolecules *33*, 2767–2771.

Rawdon, E.J. (2000). Approximating smooth thickness. J. Knot Theory Ramifications *9*, 113–145.

Rippe, K. (2001). Making contacts on a nucleic acid polymer. Trends Biochem. Sci. *26*, 733–740.

Sokal, A.D. (1996). Monte Carlo methods for the self-avoiding walk. Nucl. Phys. B. *47* (*Suppl*), 172–179.

Spector, D.L. (2003). The dynamics of chromosome organization and gene regulation. Annu. Rev. Biochem. *72*, 573–608.

Vilar, J.M., and Leibler, S. (2003). DNA looping and physical constraints on transcription regulation. J. Mol. Biol. *331*, 981–989.

Vilar, J.M., and Saiz, L. (2005). DNA looping in gene regulation: from the assembly of macromolecular complexes to the control of transcriptional noise. Curr. Opin. Genet. Dev. *15*, 136–144.

Wedemann, G., and Langowski, J. (2002). Computer simulation of the 30-nanometer chromatin fiber. Biophys. J. *82*, 2847–2859.

Woodcock, C.L., and Dimitrov, S. (2001). Higher-order structure of chromatin and chromosomes. Curr. Opin. Genet. Dev. *11*, 130–135.