BIOINFORMATICS

Vol. 15 no. 12 1999
Pages 957–964

Polymer chromosome models and Monte Carlo simulations of radiation breaking DNA

A. L. Ponomarev and R. K. Sachs

Department of Mathematics, University of California, Berkeley, CA 94720, USA

Received on February 17, 1999; revised on June 15, 1999; accepted on June 17, 1999

Abstract

Motivation: *Chromatin breakage by ionizing radiation is relevant to studies of carcinogenesis, tumor radiotherapy, biodosimetry and molecular biology. This article focuses on computer analysis of chromosome irradiation in mammalian cells.*

Methods: *Polymer physics and Monte Carlo numerical methods are used to develop a coarse-grained computational approach. Chromatin is modeled as a random walk on a cubic lattice, and the radiation tracks hitting the chromatin are modeled as straight lines hitting lattice sites. Each track can make a cluster of DSBs on a chromosome.*

Results: *The results obtained replace conjectured DNA fragment-size distribution functions in the recently developed RLC formalism by more mechanistically motivated distributions. The discrete lattice algorithm reproduces features of current radiation experiments relevant to chromatin on large scales. It approximates the continuous formalism and experimental data with adequate precision. It was also found that assuming either fixed chromatin with correlations among different clusters of DSBs or moving chromatin with no such correlations gives virtually identical numerical predictions.*

Availability: *This set of subroutines comprises the DNABreak package available at the UC Berkeley mathematical radiobiology site: http://math.berkeley.edu/~sachs/ponomarev/artem.index.html*

Contact: *alp17@math.berkeley.edu*

Abbreviations: *Mbp, megabasepair; DSB, DNA double strand break; RLC, randomly-located-clusters (Sachs* et al.*, 1999); PBC, periodic boundary conditions (Binder, 1995); RW, random walk (Sokal, 1995); size, genomic content in Mbp.*

Introduction

Recently, much emphasis has been placed on new quantitative approaches to the study of complex, fundamental biological processes, typically requiring non-traditional collaboration across disciplinary lines. Our underlying goal here is to understand chromosome geometry at very large scales during cell-cycle interphase. Computer algorithms are developed for the action of ionizing radiation on chromatin.

Chromatin

DNA is assembled with proteins to form chromatin, whose higher order structure determines large-scale chromosome geometry (Ostashevsky, 1998). It is clear that chromatin has physical and scaling properties analogous to those of long polymer chains. In fact, its molecular weight exceeds that of the largest known synthetic polymers. This fact suggests the application of the methods of polymer physics (de Gennes, 1979; Doi, 1996; Binder, 1995) to study higher order chromatin structure. These methods include analytical approaches, such as scaling (de Gennes, 1979), and numerical approaches, such as Monte Carlo algorithms that model polymer statics and dynamics (Binder, 1995; Sokal, 1995).

During interphase each chromosome is predominantly localized in a part of cell nucleus and roughly can be described as a random coil. Several sources (e.g. Ostashevsky, 1998; Yokota *et al.*, 1995) reported approximately ideal chain statistics for at least parts of chromatin in chromosomes. Ideal chains obey the following equation:

$$R_g^2 = a M, \tag{1}$$

where R_g is the average gyroradius of the molecule, which determines the average size (in nm), M is the molecular weight (proportional to genomic content), and a is a proportionality constant.

In the present studies, no excluded volume condition, corresponding to the displacement of chromatin by chromatin, is enforced, i.e. in our lattice calculations a lattice site can be occupied by more than one chromatin locus. A self-avoiding walk (Sokal, 1995) is a more complex, and perhaps more realistic, model for chromatin. However, whether chromatin is better approximated with a simple random walk (ideal chain statistics) or a self-avoiding random walk is an open fundamental question, and these properties depend not only on the structure of chromatin itself, but also on the chemistry inside the cell nucleus

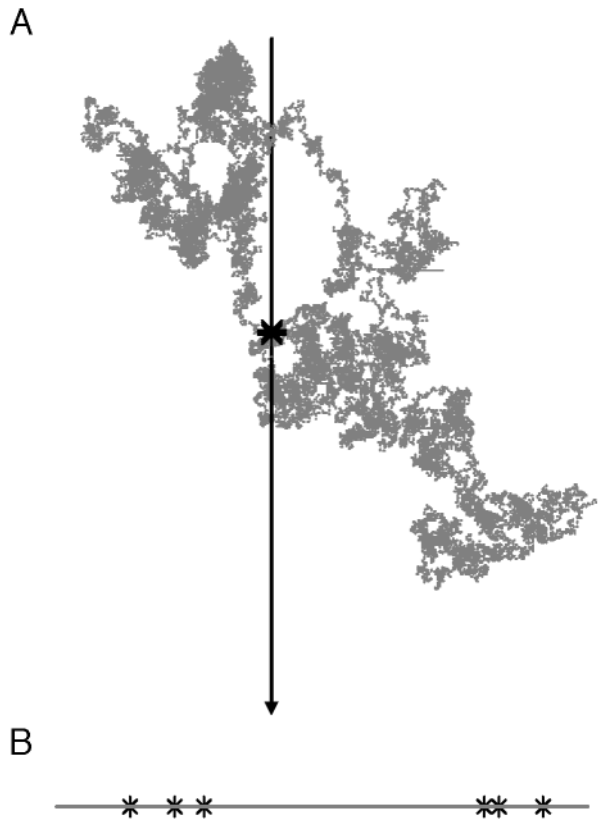

Fig. 1. DSB clustering. (A) A two-dimensional projection of an interphase chromosome, whose geometry was calculated as a three-dimensional random walk on a lattice (see text), hit by an α-particle. When a track (arrow) intersects the chromatin in space (by striking one or more of the lattice sites the chromosome occupies), it creates DSBs. (B) A small segment of the chromosome is shown magnified and straightened out, to show details of the DSB cluster, which in this particular simulation contains 6 DSBs, by indicating the relative locations of the DSBs along the chromosome.

(de Gennes, 1979; Marko and Siggia, 1997; Ostashevsky, 1998).

Another approximation made is that the same scaling property (Eq. (1)) holds for the whole chromosome except sizes less than 10 kbp. At large sizes, of the order of the whole chromosome, the Gaussian chain statistics that lead to Eq. (1) can also break down as chromosomes are often localized inside nucleus (for instance, Chinese hamster V79 or human fibroblasts chromosomes). However, for smaller chromosomes (not studied here), such as for yeast (Frankenberg *et al.*, 1999), this approximation holds true.

Radiation and DNA fragment-size distributions

The specific problem studied here is irradiation of chromatin to produce DNA double strand breaks. Chromatin inside the nucleus of a Chinese hamster cell is irradiated, as shown schematically on Figure 1a, by α-particles. Each track can produce a tight cluster of DSBs (Figure 1b). The total amount of absorbed radiation is 20–600 Gy (1 Gy = 1 J/kg). Such a dose can produce a significant number of DSBs on each chromosome. The empirical data (Newman *et al.*, 1997) were obtained by means of pulsed field gel electrophoresis (Cedervall *et al.*, 1995). During the experiment a mass spectrum of DNA fragments is found. The observed distribution of DNA fragment sizes helps characterize chromatin and track configurations (Holley and Chatterjee, 1996; Rydberg, 1996; Sachs *et al.*, 1998; Friedland *et al.*, 1998). Knowing the fragment size distributions is also important for the description of chromatin misrejoinings after irradiation, which can lead to chromosome aberrations and cellular malfunctions (Cornforth, 1998).

Radiation tracks, and their interactions with chromatin at small scales, have been subjected to very intense theoretical investigation during the last half-century (e.g. Lea, 1946; Friedland *et al.*, 1998). Less is known about large scales, in the Mbp range. The current analytic theory that predicts the large-scale distributions of DSBs on chromosomes and measured DNA fragment size distributions is the randomly-located-clusters (RLC) formalism (Sachs *et al.*, 1999). The formalism uses a one-track fragment size distribution to deduce multi-track quantities: the average fragment size, the multi-track fragment-size distribution function and the 'remaining-size' distribution function (i.e. the probability of having an interval of size s free of DSBs). This formalism has been used with a simple, conjectured one-track fragment-size distribution function, namely, the Weibull distribution (Sachs *et al.*, 1999).

Preview

The random coil picture leads to a crude, but useful numerical model, in which chromatin is represented by a discrete random walk (RW) on a cubic lattice. The probabilistic and geometric properties of RWs are well understood (Chandrasekhar, 1943; Weiss, 1994), and this provides the foundation for the DNAbreak package, based on Monte Carlo techniques. Radiation tracks will be simulated as straight lines hitting lattice points. The one-track DNA fragment-size distribution functions of the RLC formalism are thereby replaced by mechanistically derived distributions. The question of how to map physical reality onto the formal model developed here will be considered, i.e. appropriate lattice rescalings will be suggested.

We will also compute multi-track fragment-size distributions. The resulting numerical data will be compared to both the RLC formalism and experiment. It will be shown that the simulated data are consistent with the RLC formalism; that spatial correlations between clusters, neglected in the RLC formalism, play no significant role; and that adequate correspondence to the experimental results is found.

Systems and algorithm

Code

The code to model PNA double strand breaks and chromatin fragment size distribution functions was developed on UNIX and LINUX machines. The size of the package with the executable is 58 MB, and it is written in FORTRAN.

DNAbreak occupies a single directory that has the source code and help files. A Makefile allows compilation on any UNIX platform. More detailed instructions and directions are in the subdirectory called 'HOWTOs'.

Programming random walks

The simplest of the various polymer models for large-scale chromatin geometry, motivated in the Introduction, is the random walk (RW) model. It consists of simple sampling of a large number of independent configurations and calculating statistical averages of physical parameters by a Monte Carlo technique.

RW models can be continuous or discrete (Weiss, 1994), and in the limit of long chains (i.e. long walks) the two alternatives yield essentially the same results. We used a discrete model. Consider DNA loci ('monomers' in the terminology of polymer physics) along a chromosome, separated in size by some appropriate fixed amount. Successive monomers need to be spaced at least as close as the limit of resolution of the particular pulsed field gel electrophoresis experiments analyzed; yet they need to be spaced farther apart than the Kuhn length (Hahnfeldt *et al.*, 1993; Ostashevsky and Lange, 1994; Marko and Siggia, 1997) of chromatin regarded as a wormlike coil. The Kuhn length may be comparable to, or somewhat larger than, nucleosome dimensions, i.e. it may have the order of magnitude 10 kbp (Ostashevsky, 1998).

A genomic density ρ is determined by the number of monomers N in the whole chromosome (here we take $N \approx 60\,000$). The size of a 'typical' chromosome is about 280 Mbp (Chinese hamster cells). Thus a typical value of ρ was taken to be 280 Mbp/60 000 = 4.7 kbp/monomer.

Chromosome geometry is approximated by using random walk rules to place the monomers at the points of a cubic lattice, (X, Y, Z). Here X, Y, and Z are integers, and a length scale for one unit of the lattice is, very roughly, 50 nm. Note that the Kuhn segment length of chromatin is about 50 nm. Thus the chromatin is modeled on scales larger than the observed stiffness (persistence length) of chromatin. Monomers starting at one end of the chromosome are assigned integer labels, 0, 1, 2, etc. If a monomer has spatial coordinates (X, Y, Z), the next monomer then has equal probability, 1/6, of being at one of the 6 nearest neighbors, $(X+1, Y, Z), \ldots (X, Y, Z-1)$, with a random number generator determining in which of the six directions a RW grows. Each generated configuration represents chromatin at one instant. Corresponding constructions hold in one dimension or two dimensions, and are sometimes useful.

Radiation tracks

Radiation tracks were also modeled very simply. Initially it was assumed that a track goes in the Z direction and places one DSB on every monomer that occupies a corresponding lattice site (Figure 1). For example, a track might make DSBs on all monomers which have walked to (X_0, Y_0, Z), where Z is any integer, and the integer coordinates (X_0, Y_0) are the impact coordinates of the track. This construction means that, roughly speaking, the effective radius of an α-particle track is taken as the lattice spacing in the XY plane. We will discuss in the Implementation section a way to control the track size.

Following a standard PBC method (Binder, 1995), we introduce a 'fundamental square' in the (X, Y) plane of the lattice. The number of lattice sites in the fundamental square is taken as A^2. For multi-track irradiation the intensity of radiation is controlled by the parameter Λ, the average number of tracks that hit the fundamental square. Track locations are random. For example, the number of tracks hitting the fundamental square is a random variable obeying the Poisson distribution with average Λ:

$$P(n) = \exp(-\Lambda)\frac{\Lambda^n}{n!}, \tag{2}$$

where $P(n)$ is the probability of having n tracks. Similarly, the number of tracks hitting one site in the fundamental square is also Poisson, with average $\beta = \Lambda/A^2$.

Periodic boundary conditions

To avoid large Λ for a given β, the box size A is chosen to be 100–200. Thus the chromatin can be larger than the box. The PBC method (Binder, 1995) allows one to simulate an infinite radiation field by translations of tracks in the fundamental square by lA lattice units in $\pm X, \pm Y$ directions, l being an integer. As the tracks in the fundamental square are homogeneously distributed (probabilistically speaking), radiation is homogeneous everywhere. The size of the chain in lattice units is about the square root of the number of monomers ($\sim \sqrt{60\,000}$). Thus the chain covers only slightly more than one PBC box, and possible effects due to the artificial spatial periodicity of radiation are negligible.

Correlations

Two different multi-track actions are considered and compared. In the first, all the tracks hit one RW configuration. In the second, a new RW is created after each hit. The latter construction corresponds to rapidly moving chromatin, which has enough time to randomize itself between tracks.

In such a regime the chromosome relaxation time τ_R (Binder, 1995; Doi, 1996) should be less than the average time between hits. At least for small scales and some conditions (Selvin *et al.*, 1990) the chromatin relaxation times may be in millisecond range. This would be short compared to the time between two hits by different tracks on the same chromosome which is about 1 track per s at a typical radiation dose-rate 0.9 Gy/s (Newman *et al.*, 1997).

Small fragments

In the lattice computation, two or more tracks can accidentally coincide (hitting the same lattice site). If this occurs the radiation damaged monomer is assumed to have several DSBs resulting in fragments less than 1 monomer in size. Algorithmically these fragments are taken to have size 0. This is a bookkeeping device allowing one to obtain a properly normalized fragment-size distribution function.

Telomeres

There is a difficulty that exists both in theoretical and numerical treatments of the problem. A chromosome is a very long object that can often be treated as an infinite chain. However, strictly speaking it has two ends (called telomeres) that can serve as 'virtual' DSBs to produce fragments (Radivoyevitch *et al.*, 1998). Such telomere effects are somewhat analogous to surface effects in solid state physics. The code allows either taking or not taking telomere effects into account. It can be shown that these effects are insignificant here: neglecting them leads to a systematic mistake much less than the uncertainty in the empirical data.

One-track action

The DNA fragment-size distributions for one-track action are needed for benchmark calculations (see the subsection on the benchmarks below) to check the output of the DNAbreak program. They are also used in our analysis of the RLC formalism. The fragment-size distribution for one track was determined as follows. A random walk is spatially homogeneous and has no memory, so without essential loss of generality we can suppose that one end of a fragment is monomer number 0, located spatially at the origin (0, 0, 0), with the (X, Y) impact coordinates of the track therefore being (0, 0). With the model of track structure used here, and neglecting telomere effects, the next return of the RW to any lattice point of the form (0, 0, Z) defines the fragment size. It is known (from detailed properties of *two-dimensional* RWs (Weiss, 1994)) that, with probability 1, the RW will eventually return to the origin (0, 0) in the XY plane. Let g_i be the probability that the first return occurs at the jth step. Then g_j is zero unless j is even. For instance, $g_1 = 1/3$, etc; $\sum_{k=1}^{\infty} g_k = 1$.

The probability g_i was determined with a Monte Carlo simulation. A RW starting at $(0, 0, 0)$ with monomer 0 was simulated until either it returned to $(0, 0, Z)$ or a large cutoff monomer, monomer J. Repeating m times one can approximate g_i as

$$g_j = m_j/m \tag{3}$$

where m_j is the number of the first time returns at step j, and m is the number of realizations. Of interest later is the probability f_j that the first return is at size j, conditional on the return occurring on or before size J. The probability f_j and the corresponding cumulative distribution F_j are given by

$$f_j = g_j / \sum_{k=1}^{J} g_k, \quad F_j = \sum_{k=1}^{j} f_k; \quad \text{where } j = 1, 2, \ldots, J \tag{4}$$

It was found that for $J \geq 60\,000$, f_j and F_j are essentially independent of J.

Randomly-located-clusters (RLC) formalism

Dose-independent, one-track distributions can be extended to dose-dependent multi-track distributions whose clusters can overlap, complicating the fragment-size patterns. One method of extension is multi-track simulations, described above. Another one is the RLC formalism.

In the RLC formalism (Sachs *et al.*, 1999) the one-track cumulative fragment size distribution function F_j of Eq. (4) is used to deduce multi-track functions, such as the 'remaining-size' function E_j, the average fragment size and, most importantly, a cumulative multi-track fragment-size distribution B_j, where B_j is the fraction of fragments which have size less than or equal to j monomers. B_j can, but need not, take into account telomere effects.

Continuous vs discrete formalism

The RLC formalism actually uses a continuous variable s for the fragment size, and the key quantities are defined as functions of this continuous variable s : $f(s), F(s), E(s), B(s)$ etc. Here all discrete functions are matched with their continuous counterparts by taking $s = \rho j$, integrals are replaced by sums, and derivatives are given by function increments, for example, $B'(s) \leftrightarrow (B(j+1) - B(j))/\rho$.

Implementation

Benchmarks

The Monte Carlo results for the one-track distribution f_j, shown in Figure 2, were subjected to an independent check. The analytic form of f_j is well known in 1 dimension (Chandrasekhar, 1943) and can easily be found for 2 or 3 dimensions when j is small (Weiss, 1994). The numerical data were compared to these equations, thereby testing and validating the computer program.

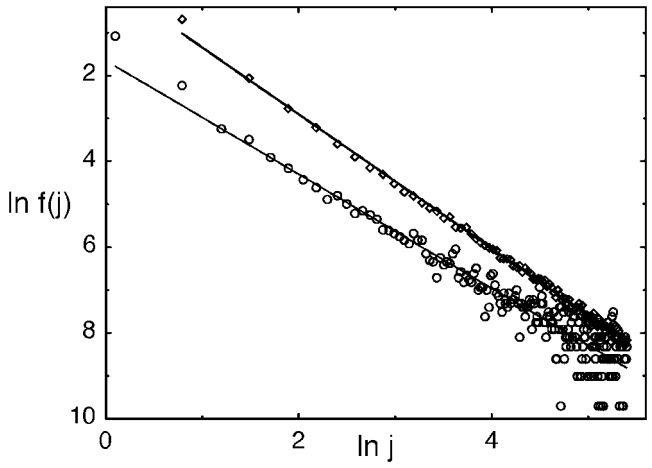

Fig. 2. Benchmarks. For software testing, the fragment-size distribution for the first return to the origin for a random walk in one dimension (diamonds) was calculated. In this case the simulations were compared to a known, simple analytic result ($f_j \sim j^{-3/2}$ for large j) (Chandrasekhar, 1943; Weiss, 1994). The Monte Carlo results are consistent with the analytic result. In higher dimensions, by using fourier series and generating functions (Weiss, 1994), the check can be done for small j. The tested program ultimately produces the one-track fragment-size distribution f_j for a 3-dimensional chromatin configuration, shown here in the log–log plot (circles). The standard error of each computer-generated point is of the order of the fluctuations seen on the figure. Lines are drawn to guide the eyes.

Rescaling

To perform simulations properly and get the right interpretation of biological results, a more systematic relation between the model and the biological parameters needs to be established besides benchmarks that confirm scaling laws. For example, let us consider radiation of Chinese hamster chromosomes (V-79 cells) by α-particles (Newman *et al.*, 1997; Sachs *et al.*, 1998). A chromosome has a certain genomic size measured in base pairs, e.g. 280 Mbp. Given radiation dose in Gy, the average number of tracks pet unit area (as well as the average energy per unit length deposited by each track) are well defined quantities. The most important model parameters are the genomic density ρ (Mbp per monomer) and the average number of tracks per lattice site β. These are input parameters. There are also calculated quantities, such as various distribution functions. The most significant is the multi-track cumulative fragment-size distribution function B_j, that allows one to calculate the mass fraction of DNA in size bins (Newman *et al.*, 1997; Sachs *et al.*, 1998). Size bins in the biological assays are genomic content intervals of DNA fragments measured in Mbp. Summing over molecular masses of fragments in these intervals yields the size fraction or mass spectrum of fragments ϕ:

$$\phi_i = \sum_{j_i}^{j_{i+1}} j\,(B_{j+1} - B_j) / \sum_{0}^{N} j\,(B_{j+1} - B_j), \qquad (5)$$

where i is the number of a bin, and $j_{i+1} - j_i = \rho^{-1}\Delta_i$, Δ_i being the genomic content interval for this bin.

The parameters of a biological assay are related to the input parameters of the model in a certain way, which need not be unique. For instance, the same chromosome can be represented either by N monomers or by $2N$ monomers. A choice of the number of monomers per chromosome introduces a genomic density parameter ρ and provides a distance scale (see above).

One notices, however, that doubling the number of monomers does not change the global geometric properties of the chromatin, the only difference occurs on the scale of a monomer. On such a small scale we disregard data, as the goal is to model global features of chromatin. We now show what input parameters (or their combinations) control the parameters of the biological assay that we have modeled.

The three following rescalings do not change the simulated data.

1. The number of monomers per chromosome N can be arbitrary for fixed β and ρ. Changing N will not change the simulated data, as long as N is not too small (≥ 2000). Note that the number of monomers N at a fixed ρ controls the genomic size of the chromosome in Mbp.

   The explanation of this fact is that the chromosome is essentially infinite in comparison with the average fragment size at high doses (equal or more than 100 Gy). The size of the chromosome defines how far apart telomeres are, and thus determines telomere effects. If the size of chromosome is effectively infinite, then telomere effects are negligible, which agrees with the numerical experiment for α-particles that showed no visible telomere effects.
2. We also found that if $\rho = \text{const}$ and $\Lambda/A^2 = \text{const}$, then the output is constant. More precisely, if A, the linear size of the fundamental square in lattice units, varies together with Λ, the number of tracks in the fundamental square, in such a way that $\beta = \Lambda/A^2$ is constant, the output parameters, namely ϕ_i, do not change. This coincides with the understanding of how PBC, simulating an infinite radiation field, should work: the size of the fundamental square is irrelevant, as long as it is not very small. This effectively eliminates the arbitrariness in choosing the PBC fundamental square size A.
3. The parameter β can be shown to be related to a distance scale, as for a given sort of radiation

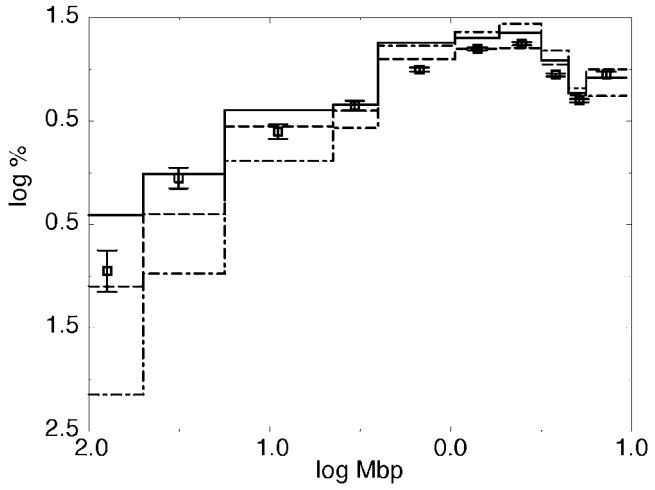

Fig. 3. Comparison of the multi-track simulation to the empirical data. The size-fraction, in percent, of DNA is shown plotted against DNA fragment-size bins in Mbp on a log–log plot. The boxes are experimental data (Newman *et al.*, 1997) for V-79 cells, irradiated with 100 Gy of He-4 ions. The best fit obtained from multi-track DNAbreak simulations, with adjustable parameters $\beta = 0.0143$ tracks per lattice site and DSB probability $p = 0.75$, is shown as the solid line. The fit is about as good as the two-parameter fit previously obtained using the RLC formalism (dashed line), based on the Weibull fragment-size distribution function chosen *ad hoc*. The dot-dashed line is the best fit based on the standard random-breakage model (Sachs *et al.*, 1998), which disregards complex cluster structure.

(≈3.5 MeV α-particles in our case) Gy units can be converted into average distance between tracks (Goodhead, 1987). Thus β defines a second scale, in addition to the size of a monomer, defined by ρ. These two scales are enough to express the biological assay parameters given by the genomic density in Mbp per monomer and the radiation density in Gy adequately in terms of the program input parameters (Λ, A and N; or ρ and β). If these two scales are changed proportionately the data (for instance, ϕ_i) should not significantly change. Numerical experiment confirmed this idea, apart from discrete-lattice size effects, relevant to small scales (discussed below). This condition can be formulated as

$$(\Lambda/A^2)/\rho = \text{const} \tag{6}$$

Track cross-section

One can control the track structure, by a parameter p, the probability to create a DSB at a given monomer hit by a track. When $p = 1$ other parameters are chosen in such a way that the apparent track cross-section (about one lattice spacing squared) is too big. This can be done by increasing ρ, but keeping $(\Lambda/A^2)/\rho = \text{const}$. Then decreasing ρ allows one to adjust the track cross-section in a least square parameter fit.

Comparison of the multi-track action to the empirical data

The least square fit based on two adjustable parameters, $(\Lambda/A^2)/\rho$ and p, results in Figure 3. The former parameter can be varied by varying Λ, for instance, and keeping A and ρ constant. The fit obtained is as good as the one based on the conjectured Weibull F_j function considered earlier in the RLC formalism (Sachs *et al.*, 1999).

Comparison of the multi-track action to the one-track action The cumulative distribution $B(s)$, obtained with the help of the RLC formalism and the one-track quantities, turned out to be the same as its discrete analogue B_j obtained *via* multi-track simulation, except for very small size bins. This confirms both the RLC formalism and the consistency of the DNAbreak package. The discrepancy on very small scales is due to the fact that the bins of small size correspond to monomer intervals 0 and 1. These small fragments were counted for bookkeeping. Lattice artifacts on this scale impose a limit on the validity of the algorithm. Another model artifact worth mentioning is that even for $p = 1$ the small size bins of the corresponding empirical data, given in Mbp, cannot be resolved in the model (i.e. these intervals can have 0 monomers).

Cluster correlation effects The RLC formalism ignores correlations between different clusters along a chromosome. A numerical experiment allows one to find the significance of such spatial correlations. In the DNAbreak formalism chromatin can be allowed to move between hits. Every successive hit is then done on the randomized chromatin. Thus the location of clusters with respect to one another along the chromosome is absolutely random. When the chromatin does not move, all hits produce DSBs on one particular chromatin configuration, which leads to some unknown correlations among clusters.

We found that these correlations are, indeed, present, but lead to only small effects. For example, let us introduce a measure of the relative deviation of B_j for the moving and fixed chromatin:

$$\nu = \frac{|B_j^{fixed} - B_j^{moving}|}{B_j^{fixed}} \tag{7}$$

For the parameters of Figure 3, ν did not exceed 1% for any j, leading to the conclusion that the cluster correlation effects can be neglected for the problem studied, as the empirical data usually have at least a 10% uncertainty.

Conclusions

The package DNAbreak allows a coarse-grained approach to interphase chromatin structure and the effects of ionizing radiation. This is a subject in which very extensive work has been done on small scales (atomic scale, or base pair scale). In the coarse-grained approach the loss of detailed information on track and chromatin structure is compensated by more realistic overall features, on the scale of a whole chromosome. The goal of such a model is to simulate the fragment-size distribution function of larger fragments, on kbp and Mbp scales. It is understood that such a model does not reproduce correct data on small scales. The model gives a satisfactory two-parameter fit of the empirical data on large genomic scales, where polymer physics techniques and Monte Carlo techniques are applicable. For smaller scales, more detailed models, such as those of Friedland *et al.* (1998) are needed.

The model proved to be consistent with the RLC formalism. The model also replaced a conjectured Weibull distribution for the one-track fragment-size function by a more mechanistically justified simulated distribution. It was found that telomere effects and spatial correlations among clusters of DSBs are negligible. A detailed analysis on what input parameters of the model control the simulated biological endpoints was made, and it was found that model has two essential parameters that correspond respectively to the track cross-section and the ratio of scales given by the radiation field and genomic density.

The present approach has a number of limitations, and various extensions would supply greater realism. Mammalian chromatin on the largest scales should perhaps be represented by a self-avoiding walk, corresponding to chromatin-chromatin interactions (Doi, 1996), and/or with loops on scales of somewhat less than 100 kbp (Friedland *et al.*, 1997), and/or with much larger loops (Sachs *et al.*, 1995; Hahnfeldt *et al.*, 1993; Ostashevsky, 1998), and/or as a constrained polymer (Dernburg *et al.*, 1996). All these extensions are technically feasible using standard polymer methods (Binder, 1995), provided enough detailed biological information becomes available. For yeast chromosomes, where relevant PFGE information has recently become available (e.g. Frankenberg *et al.*, 1999) this higher order structure is largely absent, so that comparisons to yeast would be instructive. In addition, somewhat more realistic models of radiation tracks, with a track core and a penumbra (Chatterjee and Holley, 1991) are possible even within the framework of the present coarse-grained approach, by assigning appropriate probabilities for DSB production to the impact parameter lattice point of a track in the (X, Y) plane and to neighboring lattice points.

Acknowledgements

We are grateful to D.J.Brenner, W.R.Holley and H.Paretzke for discussions. Research was supported by NIH Grant R01 GM57245.

References

Binder,K. (1995) *Monte Carlo and Molecular Dynamics Simulations in Polymer Science*. Oxford University Press, New York Oxford.

Cedervall,B., Wong,R., Albright,N. and Dynlacht,J. *et al.* (1995) Methods for the quantification of DNA double-strand breaks determined from the distribution of DNA fragment sizes measured by pulsed-field gel electrophoresis. *Radiat. Res.*, **143**, 8–16.

Chandrasekhar,S (1943) Stochastic problems in physics and astronomy. *Rev. Modern Phys.*, **15**, 1–20.

Chatterjee,A. and Holley,W.R. (1991) Energy deposition mechanisms and biochemical aspects of DNA strand breaks by ionizing radiation. *Int. J. Quantum Chem.*, **391**, 709–727.

Cornforth,M.N. (1998) Radiation-induced damage and the formation of chromosomal aberrations. In Nickoloff,J.A. and Hoekstra,M.F. (eds), *DNA Damage and Repair* Humana Press, Totowa, NJ, pp. 559–585.

Doi,M. (1996) *Introduction to Polymer Physics*. Clarendon Press, Oxford, pp. 10.

Dernburg,A.F., Broman,K.W., Fung,J.C., Marshall,W.F., Philips,J., Agard,D.A. and Sedat,J.W. (1996) Perturbation of nuclear architecture by long-distance chromosome interactions. *Cell*, **85**, 745–759.

de Gennes,P.-G. (1979) *Scaling Concepts in Polymer Physics*. Cornell University Press, Ithaca, NY, pp. 31.

Frankenberg,D., Brede,H.J., Schrewe,U.J., Steinmetz,C., Frankenberg-Schwager,M., Kasten,G. and Pralle,E. (1999) Induction of DNA double-strand breaks by 1H and 4He ions in primary human skin fibroblasts in the LET range of 8 to 124 keV/micron. *Radiat. Res.*, **151**, 540–549.

Friedland,W., Jacob,P., Paretzke,H.G., Perzl,M. and Stork,T. (1997) Simulation of strand breaks and short DNA fragments in the biophysical model PARTRAC. In Goodhead,D.T., O'Neill,P. and Menzel,H.G. (eds), *Microdosimetry: An Interdisciplinary Approach* Royal Society of Chemistry, Cambridge, pp. 43–46.

Friedland,W., Jacob,P., Paretzke,H.G. and Stork,T. (1998) Monte Carlo simulation of the production of short DNA fragments by low-linear energy transfer radiation using higher-order DNA models. *Radiat. Res.*, **150**, 170–182.

Goodhead,D.T. (1987) Relationship of microdosimetric techniques to applications in biological systems. In Kase,K., Bjarngard,H. and Attix,F. (eds), *The Dosimetry of Ionizing Radiation.* Academic Press, Orlando, pp. 1–89.

Hahnfeldt,P., Hearst,J.E., Brenner,D.J., Sachs,R.K. and Hlatky,L.R. (1993) Polymer models for interphase chromosomes. *Proc. Natl Acad. Sci. USA*, **90**, 7854–7858.

Holley,W.R. and Chatterjee,A. (1996) Clusters of DNA induced by ionizing radiation: formation of short DNA fragments. I. Theoretical modeling. *Radiat. Res.*, **145**, 188–199.

Lea,D.E. (1946) *Actions of Radiation on Living Cells*. Cambridge University Press, Cambridge.

Marko,J.F. and Siggia,E.D. (1997) Polymer models of meiotic and

mitotic chromosomes. *Mol. Cell Biol.*, **8**, 2217–2231.

Newman,H.C., Prise,K.M., Folkard,M. and Michael,B.D. (1997) DNA double-strand break distributions in X-ray and alpha-particle irradiated V79 cells: evidence for non-random breakage. *Int. J. Radiat. Biol.*, **71**, 347–363.

Ostashevsky,J.Y. (1998) A polymer model for the structural organization of chromatin loops and minibands in interphase chromosomes. *Mol. Cell Biol.*, **9**, 3031–3040.

Ostashevsky,J.Y. and Lange,C.S. (1994) The 20 nm chromatin fiber as a flexible polymer. *J. Biomolec. Struct. Dynam.*, **11**, 813–820.

Radivoyevitch,T., Hoel,D.G., Hahnfeldt,P. and Sachs,R.K. (1998) Size distributions of misrejoining DNA fragments in irradiated cells. *Math. Biosci.*, **149**, 107–136.

Rydberg,B. (1996) Clusters of DNA damage induced by ionizing radiation: formation of short DNA fragments. II. Experimental detection. *Radiat. Res.*, **145**, 200–209.

Sachs,R.K., Yokota,H., van-den-Engh,G., Trask,B. and Hearst,J.E. (1995) A random-walk/giant loop model for interphase chromosomes. *Proc. Natl Acad. Sci. USA*, **92**, 2710–2714.

Sachs,R.K., Brenner,D.J., Hahnfeldt,P.J. and Hlatky,L.R. (1998) A formalism for analyzing large-scale clustering of radiation-induced breaks along chromosomes. *Int. J Radiat. Biol.*, **74**, 185–206.

Sachs,R.K., Hahnfeldt,P.J., Ponomarev,A.L. and Hlatky,L.R. (1999) Locations of radiation-produced DNA double strand breaks along chromosomes: a stochastic cluster process formalism. *Math. Biosci.*, **159**, 165–187.

Selvin,P.R., Scalettar,B.A., Langmore,J.P., Axelrod,D., Klein,M.P. and Hearst,J.E. (1990) A polarized photobleaching study of chromatin reorientation in intact buclei. *J. Molec. Biol.*, **214**, 911–922.

Sokal, (1995) In Binder,K. (ed.), *Monte Carlo and Molecular Dynamics Simulations in Polymer Science* Oxford University Press, Oxford and New York, pp. 47–124.

Weiss,G.H. (1994) Aspects and applications of the random walk. In Stanley,H.E. and Guyon,E. (eds), *Randon Materials and Processes* North-Holland, Amsterdam, London, New York, Tokyo.

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**, 1239–1249.