Anomalous diffusion in the interphase cell nucleus: The effect of spatial correlations of chromatin

Christian C. Fritsch and Jörg Langowski

Deutsches Krebsforschungszentrum, D-69120 Heidelberg, Germany

(Received 2 February 2010; accepted 4 May 2010; published online 14 July 2010)

The metabolism of a living cell requires a permanent transfer of molecules throughout the cell and beyond its bounds. Within cell nuclei, molecules are predominantly driven by diffusion, which is influenced by the chromatin network. We propose a quantity related to the pair correlation function to measure the diffusion-relevant clumpiness of chromatin. Using Monte Carlo lattice simulations, we investigate to what extent diffusion can be anomalous due to obstruction by the chromatin network. Chromatin is simulated by a wormlike chain on a lattice, which exhibits different types of loop-induced compartmentalization on a subchromosomal level. Our simulation results show that the proposed measure of clumpiness is suitable to quantify the compartmentalization of chromatin and to connect it to diffusion anomaly parameters, critical molecule sizes for trapping and the transition lengths at which diffusion becomes normal at long times. © 2010 American Institute of Physics [doi:10.1063/1.3435345]

I. INTRODUCTION

In the cell nucleus, DNA is packed in a complex with histone proteins into nucleosomes, which are in turn compacted to a higher order structure, the chromatin fiber.\(^1\) In the interphase, the period between cell divisions, chromosomes occupy loosely distinct “territories” and form a chromatin network in the nucleus. Reading and handling the genetic information requires access and transport of a sophisticated network in the nucleus. Reading and handling the genetic information requires access and transport of a sophisticated network in the nucleus. The transport process is influenced by the distribution and the folding behavior of chromatin in a fundamental way.\(^2\)

Over distances comparable to the size of the cell, most biomolecules are transported by passive diffusion.\(^3\) If the time-dependent mean-square displacement (MSD) \(\langle r^2(t) \rangle\) of a molecule is linear in \(t\), diffusion is normal, as typically observed in bulk solutions. Diffusion is anomalous if \(\langle r^2(t) \rangle \propto t^\alpha\), with an anomaly parameter \(\alpha \neq 1\). It is called sub-diffusion if \(\alpha < 1\), which can be caused by obstruction, for instance.\(^4\)

The observation of anomalous diffusion also depends on the length scale. On a very short scale, the walker has not yet traveled far enough to encounter many obstacles by which its walk could get obstructed. The walker “sees” almost free, homogeneous space and diffusion is nearly normal. On a larger scale, the walker senses both obstacles and free space, its environment appears inhomogeneous and diffusion becomes anomalous. On a very large scale, obstruction effects average out, diffusion becomes normal again and the walker “remembers” a homogeneous medium.

Experiments using fluorescence correlation spectroscopy in cell nuclei showed subdiffusion\(^5\) and chromatin-induced obstruction\(^6,7\) of larger molecules, whereas smaller molecules diffused normally and no correlation between their mobility and the density of chromatin was found.\(^8\) The accessibility of a polymeric material to diffusing molecules primarily depends on the size of the molecules and the global density of the polymer matrix but also on the spatial arrangement of the matrix. In a homogeneous polymer network, molecules are retained if they are larger than the characteristic mesh size. At the same global density, a matrix with a strong tendency to clump together can leave space for channels that allow these molecules to travel greater distances. The spatial correlations of fluorescently labeled histones have been used to characterize the spatial arrangement of chromatin by image correlation spectroscopy.\(^9\)

Also using confocal imaging, chromatin was found to occupy about 10% of the nuclear space.\(^10\) The overall occupation of cellular compartments by macromolecules ranges from 5% up to 40% (see Ref. 11), which allows to roughly estimate an occupation of up to 30% of the cell nucleus by macromolecules other than chromatin.

In Sec. II, we describe a theoretical approach to capturing the spatial arrangement of a heterogeneous medium, such as a polymer solution, in terms of correlations and propose a measure for its “clumpiness.” For understanding the influence of the chromatin network on molecular diffusion it is important to describe the spatial arrangement of chromatin on scales below the resolution of typical light-microscopic techniques. Loop-induced compartmentalization, a concept that has been discussed for some time,\(^12-15\) influences the local chromatin density and its clumpiness on the subchromosome level. In Sec. III, we present the methods we have used to simulate interphase chromatin and diffusion of molecules on a Cartesian lattice. In order to concentrate on the purely geometric obstruction effect of chromatin, the simu-
lated chromatin is assumed immobile and diffusing molecules are affected by it only via steric volume exclusion. In Sec. IV, we present structural properties of the simulated chromatin and the results from the diffusion simulations, and we connect chromatin structure and diffusion behavior. Section V closes by discussing some implications of these results for the detection of anomalous diffusion in experiments.

II. SPATIAL CORRELATIONS OF A POLYMER

A. Excluded volume of a chain polymer with regard to a spherical molecule

We consider chromatin as a chain polymer consisting of $N$ spherical monomers with radii $R_0$ and centers $r_i$. Embedded in a solvent, this polymer occupies the volume fraction $\phi_0 = 4\pi R_0^3/3(V_{\text{vol}})^{-1}$, where $V_{\text{vol}}$ denotes the total volume of the solution. The effectively accessible space contains all possible trajectories of the diffusing molecule within the polymer solution. This space occupies less than the volume possible trajectories of the diffusing molecule within the system are equivalent. Due to intersections of the effectively accessible space is determined by the knowledge of their positions $r_i$, which are usually not available in polymer physical problems. Torquato and Stell have developed a formalism to obtain an expression for the series density $\phi_0$ of the voids since a molecule cannot be considered pointlike, and consequently some space around the molecule is excluded to the mass center of the diffusing molecule. The complement of the effectively accessible space is referred to as the excluded volume of the polymer. If there are multiple molecules diffusing at the same time, the effectively accessible space of a single molecule is further reduced. Since we are interested in the obstruction caused exclusively by chromatin, we omit crowding of the diffusing molecules in this approach.

In the lattice percolation problem, the global connectivity of clusters of free sites (pores) is determined by the universal probability $p$ of a single lattice site being unoccupied. On an infinite lattice, these clusters are isolated at low $p$, while a spanning cluster occurs sharply upon crossing the critical threshold $p_c$. The MSD of a random walker hopping between free sites gives an account of the connectivity of the unoccupied clusters. If $p < p_c$, the MSD of the walker becomes constant for long times as the walker cannot travel beyond the extent of the isolated free clusters. If $p > p_c$, the walker is able to travel arbitrarily far and it diffuses normally at long times. At the percolation threshold $p_c$, the spanning cluster is fractal on all scales, diffusion is anomalous, and the anomaly parameter $\alpha$ is specific for the type of the lattice and its dimension. Therefore, we consider the effectively accessible space as the analog to the clusters of free sites in the lattice percolation problem.

Assuming no intra- and intermolecular interactions except steric volume exclusion, the space that is excluded to a molecule with radius $R_m$ equals the union of spheres with centers $r_i$ and effective radii $R=R_0+R_m$. Both views [sketched in Fig. 1] of a spherical molecule in a polymer system are equivalent. Due to intersections of the effectively inaccessible spheres, the volume of their union is not simply given by the sum of all $N$ sphere volumes but by the alternating series.
tance \( r_{12} \) of the points and is usually referred to as the radial distribution function (RDF).

Polymer models typically do not provide correlations of orders higher than \( k = 2 \). In general, higher order correlations remain uncertain and \( \phi(R) \) can be calculated only approximately according to Eq. (3). Writing out the terms of first and second order, we get

\[
\phi(R, \theta_0) = \int \rho m(|x-r|;R)dr \cdot \frac{1}{2\pi} \int g_2(r_{12})m(|x-r_1|;R) \times m(|x-r_2|;R) dr_{12} + \cdots
\]

\[
= \phi_0(R/R_0)^3 - \frac{1}{2}(\phi_0/V_0)^2 \int g_2(a)V_2(a;R)4\pi a^2 da + \cdots,
\]

where \( V_2(a;R) \) denotes the volume of the intersection of two spheres with common radius \( R \) and center-to-center distance \( a \).

If \( \phi(R, \theta_0) - \phi_0 \ll 1 \), which is the case for small values of \( R_m = R - R_0 \), \( \phi(R, \theta_0) \) can be approximated by Eq. (5), whereas correlations of higher order are required for larger \( R_m \). Nonetheless, a strong contribution of the RDF to \( \phi(R, \theta_0) \) results in a less rapid increase with \( R_m \). Consequently, the connectedness of the effectively accessible space is related to the spatial correlations of the matrix, particularly to the RDF. This led us to define a measure of the clumpiness of chromatin using the RDF.

B. Clumpiness of the polymer matrix

While the RDF \( g_2(a) \) approaches one as \( a \to \infty \), it can in general exhibit nonmonotonicities. Therefore, it is too ambiguous to characterize its shape simply by the distance at which it drops to \( e^{-1} \) of \( g_2(0) \), which would be appropriate for an exponential decay. Instead, we introduce a more intuitive measure similar to the K-function, which has found use, e.g., in the characterization of the extent of galaxies in cosmic clusters. For a given RDF \( g_2(a) \), the expected number of sphere centers to be found within the distance \( a \) from an arbitrary sphere center is given by the K-function

\[
K(a) = \int_0^a g_2(a')4\pi a'^2 da'.
\]

As \( a \to \infty \), inhomogeneities average out and the asymptotic limit of \( K(a) \), \( (4\pi a^3/3) \), equals the expected number of sphere centers within \( a \) if the material is fully homogeneous and spatially uncorrelated. By subtracting this asymptotic from \( K(a) \), we obtain the expected excess number of sphere centers within \( a \), i.e., the contributed by clustering. The limit

\[
\lim_{a \to \infty} [K(a) - (4\pi a^3/3)] = \int_0^\infty [g_2(a) - 1]4\pi a^2 da,
\]

then gives the characteristic number of spheres that compose clusters.

Clumpiness is formally expressed by short-scale correlations. In Eq. (7), long-scale correlations contribute to \( K \) more strongly than short-scale correlations due to the factor \( 4\pi a^2 \). Therefore, as a more suitable measure for clumpiness, we define

\[
\Gamma = \int_0^\infty [g_2(a) - 1]da,
\]

and refer to it as the cumulative pair correlation of the polymer. It gives the expected excess number of clustering spheres on a beam starting in an arbitrary sphere center.

III. SIMULATION METHODS

A. Chromatin lattice model

Chromatin is simulated by a semiflexible chain polymer whose intersegment bonds are represented by sites on a Cartesian lattice. In order to account for steric volume exclusion of the polymer, distinct bonds are forbidden to occupy the same site. To maintain the connectivity of the lattice chain, distances between subsequent bonds (bond lengths) have the allowed values 1, \( \sqrt{2} \), or \( \sqrt{3} \) in lattice units. The bending stiffness of chromatin is modeled by harmonic bending potentials at each bond, giving the internal bending energy of the lattice chain

\[
\frac{E}{k_BT} = \frac{1}{2} \sum_{i=1}^{N-2} \theta_i^2,
\]

where \( N \), \( g \), and \( \theta_i \) denote the number of chain bonds, the bending rigidity constant in units of \( k_BT \), and the bond angle between bond vectors \( \vec{l}_i \) and \( \vec{l}_{i+1} \), respectively. For a wormlike chain (WLC), the average angular correlation between bond vectors \( \vec{l}_i \) and \( \vec{l}_j \) is given by

\[
\langle \vec{l}_i \cdot \vec{l}_j \rangle = \langle l \rangle^2 \exp[-|j-i|/(n_{\text{p}})],
\]

where \( \langle l \rangle \) is the average bond length and \( n_{\text{p}} \) is the persistence length, measured in units of \( \langle l \rangle \). The persistence length is the average distance along the chain at which angular correlations decay, and it can directly be related to the bending rigidity [see Appendix].

In conformations without bending rigidity, the angular correlations of bond vectors are determined by the self-avoidance of the lattice chain, and they decay at a distance shorter than the lattice constant. As a consequence of the lattice discretization, the definition of the persistence length in Eq. (10) becomes invalid, and we refer to such conformations simply as generated “without bending rigidity.”

In order to generate realizations of a canonical chromatin ensemble, we have combined the single-site bond fluctuation method for lattice polymers of Carmesin and Kremer with the Metropolis Monte Carlo (MC) method, as described earlier by Wedemeier et al. An initial chromatin conformation \( x_0 \) is iteratively altered in a sequence of MC moves. In an MC move on conformation \( x_0 \), one bond of the chain and — among its 26 neighbor sites — one trial position for its displacement is randomly picked. If the site of the trial position is already occupied by another bond, or if the chain connectivity is violated in the trial conformation, the move is rejected. Otherwise, the bending energy of the trial confor-
The lattice chain conformations representing interphase chromatin are generated in three steps: In the first, a manually generated initial conformation [Fig. 3(a)] of a chain polymer which occupies 26 sites on a $4 \times 4 \times 4$ lattice is pre-equilibrated [example of a conformation after equilibration: Fig. 3(b)].

In the second step, the lattice is rescaled by a factor of 10 in such a way that each of the previously occupied sites is being replaced by a $10 \times 10 \times 10$ cube that contains a compact chain conformation with 246 bonds [Fig. 3(c)]. There are 56 ways to connect two distinct corners of a cube, and in order to maintain the chain connectivity on the resulting $40 \times 40 \times 40$ lattice, an alignment algorithm orients these cubes accordingly. The fraction of occupied sites $26 \times 246/40^3 = 0.1$ then corresponds to the volume fraction $\phi_0$ of the unswollen chromatin chain [sketched in Fig. 1(a)] and agrees with the experimentally obtained volume occupation of the nuclear space by chromatin (see Ref. 10). A possible mechanism for the compartmentalization of chromatin is the formation of loops (e.g., see Refs. 12–15). As their origins both specific (binding) and unspecific (entropic) forces have been discussed. In the model presented here we have implemented a specific formation: loops are generated during the further equilibration of the conformation obtained so far by retaining the adjacency of each two bonds that compose a loop vertex.

The third step of our nested equilibration procedure is performed while the following individual loop formation constraints hold. The WLC type [sketched in Fig. 2(e)] is produced without loop constraints and represents the WLC. The giant loop (GL) type [Fig. 2(d)] features loops with lengths above the order of the typical persistence length of chromatin (GL model for chromatin13). In the loose loops (LL) type [Fig. 2(c)], loops are loosely aligned along the chain, whereas 5 loops in the 5-loop subcompartment (LS) type [Fig. 2(b)] and 10 loops in the 10-LS type [Fig. 2(a)] again fold into a GL consisting of 220 bonds, thus forming a multiloop subcompartment.13 The lengths of the small loops in the latter three types are of the order of the persistence length (22 bonds in types LL and 10-LS, 44 bonds in type 5-LS). For the sake of clarity the parameters for the loop formation of the five folding types are listed in Table I. In Figs. 3(d)–3(f), typical conformations after the third equilibration step are shown.

Since the extent of our simulation lattice is just about 10% that of a real cell nucleus, the use of hard lattice boundaries would cause a confinement of chromatin and the diffusing molecules to an unrealistically small portion of space. This technical limitation is avoided by allowing periodic boundary conditions (PBC) in all directions of the lattice, and the simulated space becomes virtually infinite. As a consequence of that, however, conclusions on the behavior of chromatin and the diffusing molecules upon a scale that cor-
responds to the real extent of a nucleus are invalid; the confinement by the real boundaries is not accounted for in this way. The PBC in the simulations enter at the very first step of the nested equilibration procedure where the manually created initial configuration is equilibrated [step between stages illustrated in Figs. 3(a) and 3(b)].

For the conformation on the \(4 \times 4 \times 40\) lattice we assume a lattice constant \(a_{40}=30\) nm, which corresponds to the diameter of the chromatin fiber. With an average bond length \(\langle l \rangle \approx 1.44\), the simulated conformational states have \(280 \mu\text{m}\) length, which is about one quarter of the human chromosome 21.

The equilibration on the \(4 \times 4 \times 40\) lattice is performed in \(10^7\) steps with a bending rigidity equivalent to a persistence length \(L_p=2\langle l \rangle\), which ensures a good unfolding of the initial conformation. For each of the five folding types, the following equilibration on the \(4 \times 4 \times 40\) lattice is first performed in \(10^8\) steps with persistence length \(L_p=2\langle l \rangle\) for fast disentanglement. During further \(10^8\) steps, the resulting conformations are then again equilibrated separately without bending rigidity or with six different persistence lengths between \(L_p=1\langle l \rangle\) and \(L_p=4\langle l \rangle\), which correspond to real lengths between about 45 nm and 180 nm. 50 independent realizations are generated for each folding type and persistence length.

C. Diffusion simulation

In the environment of a single-site lattice chain polymer, only molecules of at least the same thickness as the polymer can be represented. Diameters of diffusing proteins in the cell nucleus come down to a few nanometers, and in order to provide for these smaller molecules, the lattice is further re-sized by a factor of 3, yielding a chain environment composed of \(3 \times 3 \times 3\) cubes on a \(120 \times 120 \times 120\) lattice.

Diffusion of molecules in the chromatin environment is simulated by random walks with first neighbor displacements. A walker of particular size is represented by a quasi-spherical body that occupies the \(N\) nearest sites within a specific radius \(R_{\text{max}}(N)\) of a center site. The walker radius \(R_m\) is then calculated as a smooth sphere radius such that \(N=4\pi R_m^3/3\).

According to the considerations in Sec. II, instead of displacing all \(N\) sites of a lattice sphere for a random walk step, only its center site is displaced. The center site can, however, access only those sites of the lattice that belong to the effectively accessible space of the walker. The excluded volume of the chromatin chain is generated by swelling up the chain by all the lattice sites within the particular distance \(R_{\text{max}}\) of each of its sites.

The MSD is recorded as a function of time \(t\) (measured in MC steps) in every walk. If an attempted walker displacement in the occupation of a site by both the walker and the excluded volume of the lattice chromatin, the displacement is not executed, but the time is increased by 1. The bonds of the chromatin chain are not fluctuating during a walk so that the walker senses chromatin as an immobile obstacle.

For each chromatin conformation we simulated random walks of lattice spheres with 60 different sizes. A walk consisted of \(10^7\) walker displacements and its MSD was recorded as the average over \(10^3\) walks within a single chromatin conformation sample. This MSD was again averaged over the 50 conformation samples. Hence we obtained the ensemble average of the MSD, which may deviate essentially from the time average in case of subdiffusion. In particular, the time average may suggest normal diffusion although the actual diffusion process is in fact subdiffusive. In a disordered system such as the cell nucleus, this discrepancy can arise from its inhomogeneity in such a way that single trajectories may differ greatly from each other. This type of ergodicity breaking is particularly relevant when probing the movement of molecules in single particle tracking experiments.

IV. RESULTS

A. Chromatin as a semiflexible chain polymer

We investigated the influence of local folding of the chromatin chain on its global properties. At a given persistence length, conformational states with loops are generally more bent and cause an increase of the mean bending energy per bond. The average bond length \(\langle l \rangle\) in lattice units (l.u.) is between 1.42 in conformations without bending rigidity and 1.46 in conformations with high bending rigidity. The internal bending rigidity dominates the topology at short contour lengths.

The angular correlations of bond vectors show the expected exponential decay according to Eq. (10) only in WLC type conformations [see Fig. 4(a)]. In the case of long loops (GL type), the angular correlations deviate slightly from the WLC model; the deviations are more pronounced in the case of short loops (LL, 5-LS, and 10-LS types). Within a loop, the bond vectors at bonds separated by half a loop length along the contour of the chain are mostly antipodal, causing angular anticorrelation.

The WLC model implies a mean-square end-to-end distance

\[
\langle R^2 \rangle = 2 L_p \ell - 2 L_p^2 \left[ 1 - \exp(-\ell/L_p) \right],
\]

for contour length \(\ell\). This is in agreement with the relation of \(\langle R^2 \rangle\) and \(\ell\) of chromatin WLC type conformations at small and intermediate \(\ell\) [see Fig. 4(b)]. The initial chromatin conformations [Fig. 3(c)] were constructed in a rosettelike arrangement, similar to metaphase chromosomes. Equation
First, we obtained the matrix in chromatin conformations from our lattice simulations. These maps confirmed the presence of chromosome territories: chromosomes were statistically processed such that they all crosslinking probabilities between genomic sites on human bonds. To reduce the amount of data, we calculated a tenfold smaller corresponding 10 interaction coefficient of the entry of a section of chromatin with a contour are also spatially close. It is evident that the 10-LS type features a conceivable globule intermediate contour lengths corresponding to multiloop sub-compartments, the WLC model with low bending rigidity less so, and the WLC model with Lp=90 nm is least correlated locally. This can be rationalized by a tendency of a stiffer chain to fold through space over larger distances. In order for the chain to fill the whole nucleus, many distant parts will then come into closer contact. On the other hand, high flexibility and, even more so, local loop constraints favor contacts between segments that are already close together on the linear chain sequence.

For a statistical treatment of the different chromatin folding type ensembles, we calculated the intramolecular RDF of a polymer, g2,intra. It is proportional to the probability of finding a monomer of the same polymer within some radial distance, and it shows the fractal nature of the polymer. The fractal dimension dI is related to the intramolecular RDF by g2,intra(r)~r^{dI-3}. WLC type conformations which have been generated without internal bending rigidity behave like ideal chains (dI=2) on shorter distances up to r=540 nm [Fig. 6(a)]. Above that, the excluded volume interaction dominates the folding of the chains, and they behave like

\[
\langle R \rangle \propto \ell^{1/3}
\]

which deviates at large \( \ell \) because this compact structure is retained to a certain extent due to the self-obstruction of chromatin. It was argued earlier that chromosome territories actually form due to this effect, and that interphase nuclei never equilibrate. In our simulations, an agreement with the WLC model on all scales would require full equilibration, which costs too much computation time if performed with the whole set of parameters. However, long-range correlations do not influence local diffusion behavior, and therefore we restricted the generation of chromatin to this moderate equilibration. For conformations with loops, \( \langle R^2 \rangle \) levels off more rapidly with increasing \( \ell \) compared to the WLC. At intermediate contour lengths corresponding to multiloop sub-compartments, the 10-LS type features a conceivable globule polymer regime, in which \( \langle R \rangle \approx \ell^{1/3} \) holds.

**B. Spatial correlations**

Different folding constraints of chromatin lead to different spatial arrangements. The third row of Fig. 3 illustrates the emergence of chromatin territories induced by loop formation constraints: Fig. 3(g) shows a 10-LS type chromatin conformation generated without bending rigidity. The spatial variation of the color indicating the position on the contour is small and the occupation of the lattice is inhomogeneous. The WLC type conformation without bending rigidity in Fig. 3(h) is more intermingled and distributed more homogeneously. This effect is even more pronounced with higher bending rigidity [Fig. 3(i)]

The three-dimensional folding of a polymer determines the spatial correlations of its monomers. In turn, the spatial correlations between specific sites on a chromosome give an account of its folding within the cell nucleus. This has recently been exploited by several large-scale genomic interaction analyses using chromatin crosslinking and deep sequencing. In Ref. 15, the experimentally obtained crosslinking probabilities between genomic sites on human chromosomes were statistically processed such that they allowed to construct proximity maps of the human genome. These maps confirmed the presence of chromosome territories.

Similar to that, we constructed proximity maps of sites in chromatin conformations from our lattice simulations. First, we obtained the matrix \( R^* \) of distances \( r_{ij}^* \) between bonds \( i \) and \( j \). \( R^* \) is a very large matrix, and in order to reduce the amount of data, we calculated a tenfold smaller matrix \( R \) in which \( r_{ij} \) is the average value of the entries in the corresponding 10x10 submatrix of \( R^* \). Then we obtained the correlation matrix \( C \) in which \( c_{ij} \) is the Pearson correlation coefficient of the \( i \)th row and the \( j \)th column of \( R \). Every entry of \( C \) corresponds to a section of chromatin with a contour length of about 440 nm. The \( c_{ij} \) take values between -1 (distant pair of sections) and +1 (close pair of sections). In Fig. 5, the correlation matrices \( C \) of the three conformations that are shown in the second row of Fig. 3 are displayed. Large yellow blocks indicate compartmentalization: a large number of chromatin sections that are close on the chromatin contour are also spatially close. It is evident that the 10-LS model has the strongest tendency to fold into spatially separated subcompartments, the WLC model with low bending rigidity less so, and the WLC model with \( L_p=90 \) nm is least correlated locally. This can be rationalized by a tendency of a stiffer chain to fold through space over larger distances. In order for the chain to fill the whole nucleus, many distant parts will then come into closer contact. On the other hand, high flexibility and, even more so, local loop constraints favor contacts between segments that are already close together on the linear chain sequence.

![Fig. 5](image_url)  
**Fig. 5.** Spatial correlation of chromatin sections displayed as proximity maps. Panels [(a)-(c)] show the correlation matrices of a 10-LS type without bending rigidity, a WLC type without bending rigidity and a WLC type with \( L_p=90 \) nm, respectively. Blue: correlation coefficient -1, yellow: correlation coefficient +1.

![Fig. 6](image_url)  
**Fig. 6.** Spatial correlations of chromatin. (a) Intramolecular RDF \( g_2,intra \). (b) RDF \( g_2 \) of the solution. Dotted line: scaling of an ideal chain (\( r^{-3} \)), dashed-dotted line: scaling of a self-avoiding walk (\( r^{-4/3} \)), solid line: semifluid polymer solution model. Red squares: WLC type, blue circles: GL type, green diamonds: 10-LS type.
self-avoiding walks \((d_\text{c}=5/3)\). The formation of loops clearly induces multifractality: Up to very short distances around \(r\approx150\ \text{nm}\), conformations with loops are more compact than WLCs, and \(g_{2,\text{intra}}(r)\) scales with an exponent greater than \(-1\). At larger \(r\), \(g_{2,\text{intra}}(r)\) scales with an exponent even less than \(-4/3\) of self-avoiding walks. On the global scale, which is given by the size of the underlying lattice, all folding types have the same density.

A molecule diffusing in a polymer solution is obstructed equally by all polymers. The intramolecular correlations of a single polymer in a solution of many polymers, however, do not give a full account of the concentration of all obstructing polymers; the intermolecular correlations are missing. Therefore, knowledge of the folding of a single chromosome is not sufficient if intermingling with other chromosomes cannot be entirely excluded. This makes \(g_{2,\text{intra}}\) useless for a connection to the diffusion behavior.

This is where the RDF \(g_2\) defined in Sec. II comes into play. It contains both intramolecular and intermolecular correlations. For chain polymers in semidilute solution it is given by\(^{36}\)

\[
g_2(r) = \frac{3}{c} \pi b^2 r \exp(-r/\xi) + 1,
\]

where \(\xi, c,\) and \(b\) are the correlation length of the polymer, its concentration, and the effective bond length, respectively. Only for conformations of WLC and GL type and apart from deviations due to lattice discretization for small \(r\), \(g_2\) is well described by Eq. (12) [see Fig. 6(b)]. The inserted parameters are \(\xi=2, c=0.1,\) and \(b=1.8\). The clumpier conformations with small loops (LL, 5-LS, and 10-LS types) do not fit this model and invalidate \(\xi\) as a measure for spatial correlations of the solution. Nonetheless, \(g_2\) contains essential information for a diffusion-relevant quantification of chromatin folding, and we employed the cumulative pair correlation \(I\) from Eq. (8) to connect this information to diffusion.

**C. Data from diffusion simulations**

During the diffusion simulations the MSD of a diffusing molecule is recorded as a function of time \(t\) (measured in MC steps). The MSD curves are used to calculate diffusion coefficients \(D\), anomaly parameters \(\alpha\) and critical particle radii \(R_c\). At each time \(t\), a function proportional to \(t^\alpha\) is fitted to the ten subsequent data points. Depending on \(t\), this function gives different values of \(\alpha\), of which the minimal and the maximal values, \(\alpha_{\text{min}}\) and \(\alpha_{\text{max}}\), are calculated [illustrated in Fig. 7]. In the long time regime, \(\alpha\) approaches either 1 if the effectively accessible volume percolates, or 0 if the diffusing molecule is confined within a finite region.

The transition time \(t_\alpha\) from anomalous to normal diffusion can be obtained by intersecting the asymptotics of the intermediate and the long time regimes (regimes where \(\alpha_{\text{min}}\) and \(\alpha_{\text{max}}\) are determined). Let \(d_{\text{tr}}(t_\alpha)\) denote the corresponding root-mean-square displacement. In the case \(\alpha_{\text{min}}=0\), diffusion is anomalous at all times \(t\), and \(d_{\text{tr}}\) signifies the characteristic distance a molecule can travel within the region it is confined in.

Unlike the spatial dimensions in the diffusion simulations (lattice constant \(a_{120}=10\ \text{nm}\)), the time \(t\) (and therefore the diffusion coefficient \(D\)) still lacks a scaling factor. In an obstacle-free environment, the step length per time step equals 1 for any molecule radius \(R_m\) at any time \(t\). Consequently, the diffusion coefficient is \(D_0=2\pi\eta a^2\) at \(T=300\ \text{K}\), which would imply \(\alpha_{\text{fr}}=0\) at

\[
D = \frac{2\pi\eta a^2}{k_B T} (6\pi\eta a_{120} R_m)^{-1} = 2.1 \times 10^{-11} \frac{\text{m}^2}{\text{s}}\times R_m^{-1}.
\]

**D. Percolation threshold of the effectively accessible space**

Under what circumstances is a molecule trapped in the chromatin network? The confinement of a diffusing molecule depends on the connectivity of the effectively accessible space. We have calculated the excluded volume fraction \(\phi\) as described in Sec. III C for every conformation and molecule size. The dependences of the diffusion coefficient \(D\) and the anomaly parameter \(\alpha_{\text{min}}\) of \(\phi\) are shown in Fig. 8. \(D(\phi)\) and \(\alpha_{\text{min}}(\phi)\) lie close to but not precisely on master curves. In particular, the universal relation \(D=1-\phi/\phi_c\) presumed earlier\(^{18,37}\) does not hold. The site percolation threshold of a three-dimensional first-neighbor Cartesian lattice is \(\phi_c=0.311\ 605\), which would imply \(D=0\) at

\[
\phi_c=0.311\ 605,
\]

\[
\phi=0.2
\]

\[
\phi=0.8
\]

\[
\phi=1
\]

\[
\phi=0.1
\]

\[
\phi=0.5
\]

\[
\phi=0.9
\]

\[
\phi=0.99
\]

\[
\phi=1
\]

\[
\phi=0.1
\]

\[
\phi=0.5
\]

\[
\phi=0.9
\]

\[
\phi=0.99
\]

\[
\phi=1
\]

\[
\phi=0.1
\]

\[
\phi=0.5
\]

\[
\phi=0.9
\]

\[
\phi=0.99
\]

\[
\phi=1
\]

\[
\phi=0.1
\]

\[
\phi=0.5
\]

\[
\phi=0.9
\]

\[
\phi=0.99
\]

\[
\phi=1
\]
reduced diffusion coefficient is given by $\phi_c = 1 - 0.311605 = 0.688395$. We obtained values of $\phi_c$ in the range between 0.88 and 0.98, depending on the underlying chromatin structure. [Fig. 8(a)]. Although the results cannot clearly disprove the linearity in $D(\phi)$, they question the application of this relation to polymer systems.

### E. Connecting chromatin structure and molecular diffusion

The dependence of the diffusion behavior upon the size of a molecule varies with the underlying chromatin structure. The data examples for diffusion coefficients and anomaly parameters in Figs. 9(a) and 9(b) indicate the tendency that compartmentalization leads to a less steep decrease of $\alpha_{\text{min}}$ and $D$ with rising $R_m$. In Amsden’s diffusion model,\(^{39}\) the reduced diffusion coefficient is given by

\[
\frac{D_{\text{ph}}}{D_{\text{ph,0}}} = \exp \left[ -\pi \frac{R_m + R_0}{R_0} \frac{\phi_c}{(k + 2\phi_c^0)^2} \right],
\]

where $k$ is related to the persistence length of the polymer. While this model fits the diffusion coefficients in the WLC type [solid line on red squares in Fig. 9(a)], it generally fails for chromatin with loops.

The cumulative pair correlation $\Gamma$ is useful to relate macroscopically measurable quantities, for instance, the anomaly parameter $\alpha_{\text{min}}$, to the clumpiness of the network. Figure 10(a) shows that $\alpha_{\text{min}}$ decreases monotonously with $\Gamma$. Thus, the larger $\Gamma$ is, the larger $R_m$ can be at given $\alpha_{\text{min}}$. This also holds for $D$ (not shown).

The critical radius $R_c$ of a molecule is the radius at which the effectively accessible space no longer percolates, i.e., the molecule is trapped. Qualitatively, it manifests itself in the transition from the upper curve to the lower curve shown in Fig. 7. We measure $R_c$ as the $R_m$ at which $\alpha_{\text{min}}$ (or $D$) drops to 0. Assuming, e.g., chromatin with a persistence length 130 nm and a volume fraction of 0.1, a Cajal body of 200 nm diameter gets trapped in a WLC environment (WLC type), while it can travel arbitrarily far with $D_{\text{ph}} = 2 \times 10^{-14}$ m² s⁻¹ in an environment with ten-loop subcompartments (10-LS type). Our simulation data suggest a linear relation $R_c = (\Gamma - 0.6) \times 80$ nm [see Fig. 10(b)].

To illustrate the difference between confined and unconfined diffusion, representative trajectories from the simulations are shown in Fig. 11. The corresponding squared displacements are displayed in Fig. 12.

### F. Transition from anomalous to normal diffusion

Different chromatin folding types will show different cumulative pair correlations. Therefore, the structural inhomogeneities will also average out on different length scales. Consequently, the length scales for the transition from anomalous to normal diffusion are also different. The transition distance $d_{tr}$ in, e.g., the 10-LS type is shifted to larger values compared to the WLC type [see Fig. 13(a)]. According to this, the transition times $t_{tr}$ of the 10-LS type exceed that of the WLC type, but only at small radii $R_m$. At larger
V. CONCLUSION

We have studied the influence of crowding by interphase chromatin on the diffusion of spherical molecules in the cell nucleus. In particular, we asked how an inhomogeneous distribution and the folding topology of the chromatin chain influences intranuclear transport. Chromatin with different types of subchromosomal folding was simulated by a chain polymer on a Cartesian lattice. The applicability of this discrete model had been confirmed earlier in a comparison to a continuous model.\textsuperscript{18,40}

It is evident that those chromatin chains that facilitate local folding (looped conformations, WLC with short $L_p$) also promote compartmentalization [see third row of Figs. 3 and 5]. Stiffer and locally unconstrained chains curve back on themselves at longer contour lengths, and the crumpling of the chain into compartments is suppressed.

Small molecules with diameters up to about 20 nm show the same diffusion coefficients $D$ (or rather $D_{ph}/D_{ph,0}$) and anomaly parameters $\alpha_{min}$ in all types of chromatin networks [see Fig. 9]. We therefore conclude that the particular folding topology of chromatin has no effect on the diffusion of small molecules. However, larger molecules are strongly influenced by the folding topology. Their diffusion is less obstructed in those networks with high compartmentalization.

We have proposed a method to characterize the network clumpiness that goes along with the compartmentalization of chromatin. The cumulative pair correlation $\Gamma$, which can be assigned to a network with a particular folding topology, is suitable to estimate the critical radii $R_c$ at which diffusing molecules get trapped [see Fig. 10(b)]. In the presence of chromatin network dynamics such molecules can travel even further.\textsuperscript{41} $\Gamma$ can therefore be used as an indicator for the requirement of network dynamics.

In many approaches, the polymer volume fraction $\phi_0$ has been connected to diffusion coefficients of small molecules in polymeric materials (e.g., for a review see Ref. 42). $\phi_0$ loses its significance as a global quantity if the material is strongly inhomogeneous on a local scale. This can be the case if the polymer is forced to form subcompartments, and then the diffusion behavior on a global scale is rather connected to the excluded volume fraction $\phi$. In Sec. II, we demonstrated how $\phi$ is determined by spatial correlation functions, which account for inhomogeneity on all scales.

Several earlier studies showed that the percolation threshold of the effectively accessible volume fraction of a system with inhomogeneously distributed obstacles is lower than that of a system with homogeneously distributed obstacles (e.g., see Refs. 43–45). Our simulations confirm this effect for systems of long polymer chains. Consequently, uncorrelated arrangements of obstacles are inappropriate to model critical transport phenomena in macromolecular networks such as chromatin.
A diffusion coefficient or an anomaly parameter is not unique to a molecule of given size in a particular environment. Our simulations underline the fact that both also depend on the length scale of observation. The transition from anomalous to normal diffusion implies that the observed diffusion is anomalous whenever molecules are traced within time intervals shorter than \( t_u \) or, equivalently, on length scales shorter than \( d_u \). Many studies have been made on intracellular mobility using fluorescence correlation spectroscopy.\(^{5,6,40,47}\) There, the motion of fluorescent probes is measured in a laser focus. The maximal root-mean-square displacement that can be covered within such a focus is about 1 \( \mu \text{m} \). Referring to Fig. 13(a), all values of \( d_u \) are below the typical focus diameter. This indicates that the transition to normal diffusion can be observed in all cases of molecules with radius \( R_m < R_u \), provided the time resolution of the experimental setup is sufficient.

Clumpiness shifts the transition distances \( d_u \) in our chain systems to values larger than expected for random systems. If subdiffusion originates not only from obstruction by chromatin but also from other effects such as crowding of the diffusing molecules or temporary sticking to chromatin, their contribution to anomaly is overestimated if randomness of the matrix is assumed. A comparison of the effects that interfere with the obstruction by chromatin will be made elsewhere.

Furthermore, our chromatin model is suitable for the simulation of whole chromosomes based on experimentally obtained crosslinking probabilities of genomic sites. These probabilities or, similarly, the spatial correlations between genomic sites [as displayed on the proximity maps in Fig. 5] carry information which can be incorporated into the presented Metropolis MC procedure. Resulting chromatin conformations then exhibit the desired (experimentally obtained) spatial correlations and provide information about the so far insufficiently known three-dimensional arrangement of chromosomes.

### ACKNOWLEDGMENTS

C.C.F. thanks M. Niezgódka and A. Trykozko for their hospitality and valuable discussions during research stays at the ICM, University of Warsaw. C.C.F. is supported by the International Graduiertenkolleg IGK 710 at the IWR, University of Heidelberg.

### APPENDIX: PERSISTENCE LENGTH OF A WORMLIKE CHAIN ON A CARTESIAN LATTICE

For chain polymers with fixed bond length \( l \), fixed bond angle \( \theta \) and free bond torsion, the component of the \((i+1)\)th bond vector along the \( i \)th bond vector is \( l \cos \theta \). Consequently, the average angular correlation between arbitrary bond vectors \( \vec{l}_i \) and \( \vec{l}_j \) is \( \langle \vec{l}_i \cdot \vec{l}_j \rangle = l^2 \langle \cos \theta \rangle |i-j| \). For WLCs, this holds only for average bond lengths and bond angles, and we write

\[
\langle \vec{l}_i \cdot \vec{l}_j \rangle = \langle l \rangle^2 \langle \cos \theta \rangle |i-j| = \langle l \rangle^2 \exp[|j-i| \ln \langle \cos \theta \rangle],
\]

(1A)

from which we obtain Eq. (10) when we define the persistence length \( L_p = \langle l \rangle / \langle \cos \theta \rangle \).

The average cosine of the bending angle is related to the bending rigidity \( g \) by\(^{48}\)

\[
\langle \cos \theta \rangle = \frac{\int_0^\infty \cos \theta \sin \theta \exp[-g \theta^2 / 2k_BT] \, d\theta}{\int_0^\infty \sin \theta \exp[-g \theta^2 / 2k_BT] \, d\theta}. \tag{2A}
\]

On the lattice, the integration in Eq. (2A) is to be replaced by a summation over all possible bond angles \( \theta \). We require their statistical weights \( P(\theta) \) to transform Eq. (2A) into

\[
\langle \cos \theta \rangle = \frac{\sum \rho(\theta) \cos \theta P(\theta) \exp[-g \theta^2 / 2k_BT]}{\sum \rho(\theta) \exp[-g \theta^2 / 2k_BT]} \tag{3A}
\]

For that, we assume first that we are given a particular bond vector \( \vec{l}_i \). In order to obtain the bond angle \( \theta \), there are \( N(\theta; l_1, l_2) \) ways to select a second bond vector \( \vec{l}_2 \). Furthermore, for each of these choices of \( \vec{l}_2 \), there is a number of ways to rotate both vectors at once on the lattice. This number depends on the length of the shorter of the two bond vectors, for instance, on \( l_1 \), and we denote it by \( N(l_1) \).

Hence, the statistical weight of a realization of \( \theta \) by the bond vectors \( \vec{l}_i \) and \( \vec{l}_j \) is \( P(\theta) = N(\theta; l_1, l_2) N(l_1) / \sum \rho N(\theta; l_1, l_2) N(l_1) \). All occurring cosines, the according \( P(\theta; l_1, l_2) \) and the \( P(l_1) \) are shown in Table II. Finally, combining Eqs. (1A) and (3A) allows for calculating the persistence length \( L_p \) as a function of the bending rigidity \( g \).

\begin{table}[h]
\begin{center}
\small
\begin{tabular}{|c|c|c|c|c|c|}
\hline
\( (l_1, l_2) \) & (1, 1) & (1, 2) & (1, 3) & (2, 2) & (2, 3) & (3, 3) \\
\hline
\cos \theta & 1, 0 & 1, -1 & 1, -1 & 1, 1/2 & 1, -1/2 & 1, 1/3 & 1, -1/3 \\
\hline
\( N(\theta; l_1, l_2) \) & 1, 4 & 4, 4 & 4, 4 & 4, 2 & 4, 2 & 1, 3 & 3 \\
\hline
\( N(l_1) \) & 6 & 12 & 8 & & & & \\
\hline
\end{tabular}
\end{center}
\end{table}


J. Langowski, Nucleus 1, 37 (2010).


