Correlations in systems of complex directed macromolecules

Ch Moseley and K Ziegler

Institut für Physik, Universität Augsburg, Germany

Abstract

An ensemble of directed macromolecules on a lattice is considered, where the constituting molecules are chosen randomly with N different colours. Molecules of the same colour experience a hard-core (exclusion) interaction. We study the robustness of the macromolecules with respect to breaking and changing the colour of a constituting molecule, using a $1/N$ expansion. The properties depend strongly on the density of macromolecules. In particular, the macromolecules are robust against breaking and changing the colour at high densities but fragile at intermediate densities.

1. Introduction

In a liquid of molecules long chain-like macromolecules (MM) can be formed due to a special steric intermolecular interaction. These MM are often called polymers [1]. We assume that the MM are directed, i.e. they have a preferred direction in the d-dimensional space. This implies that they are stretched and do not form loops or overhangs.

In many systems the constituting molecules (CM) of the MM are identical. However, in more complex systems, like in biological substances, the CM are chemically different. An example for the latter are the sequences of amino acids in proteins [2–4] and molecular sequences along a DNA helix. For our statistical model we assume that there are N different kinds of molecules, represented by N colours. They are randomly chosen to contribute to the MM such that the resulting MM can be viewed as a chain of coloured beads (cf figure 1). An ensemble of MM can be considered as being created in a liquid of molecules by a steric interaction among the CM. As a result, we have a space-filling system of molecules, where some are constituting the MM and others exist as individual molecules. Properties of the ensemble on large scales should not depend on local structures. This allows us to consider a lattice model [3], where each lattice site is occupied by a molecule of each type. It is assumed that molecules of the same type have a strong repulsive interaction. This restricts us to at most one molecule of each type at any given lattice site. The fact that the interaction of different molecule types is neglected can be understood as a chemical property of the molecules.
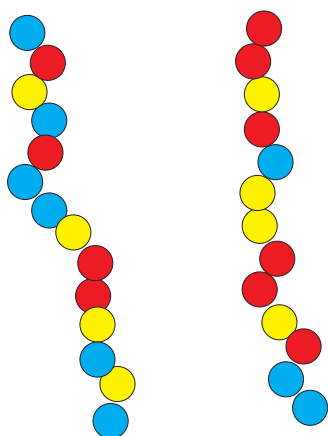
Figure 1. Two directed macromolecules.

Any pair of neighbouring molecules can form a bond state, representing a pair of molecules, which acts as a building block for an MM. There is a tendency of these bond states to bond with other molecules. This leads eventually to an equilibrium state, where we have a mixture of only large MM and individual molecules. The repulsive interaction between the molecules creates correlations among the MM which lead to interesting properties of the ensemble. For instance, the density of MM as a function of the probability of a local bond state is of interest, as well as the question of how likely it is to break an MM or to replace a CM of special type by one of a different type. The latter can happen during a collision process when an individual molecule collides with an MM and exchanges with a CM of the MM.

We consider an equilibrium ensemble of lattice MM and define three quantities to describe its global properties with respect to the molecular structure. There are two quantities that measure the correlations in real space. One measures how likely it is to break a directed MM and take it apart in space (cf figure 2); the other one measures how likely it is to create a new (finite) MM by changing the type of the CM in an MM. If the ends of the broken MM or the ends of the new (finite) MM are x and x' we write for the corresponding quantities $G_{xx'}$ and $H_{xx'}$, respectively. It will be shown that $G_{xx'}$ and $H_{xx'}$ are directly related. Another quantity is the coherence in α (molecular-component) space. This is called $S_{\alpha\alpha'}$ and measures how likely it is to replace a CM of type α in an MM by one of type α' (cf figure 3). All these quantities are evaluated in the limit $N \to \infty$ and in a $1/N$ expansion. We apply a functional-integral formalism developed in [5] to perform the calculations in a straightforward manner.

2. The model

A statistical model for the equilibrium statistics of directed MM was introduced in [5]. We briefly summarize the construction of the model and then turn to the evaluation of the interesting correlations mentioned in the introduction. The model is based on the idea that on each site of a (cubic) lattice N different molecules are located. They exist either as individual (separated) molecules of a liquid with relative weight μ or contribute to a directed MM as a bond state of a nearest-neighbour molecular pair at x, x' of types α and α'. Moreover, the MM are directed, i.e. there is a special direction t in the d-dimensional space as a preferred direction of the MM. The fluctuations of the MM are perpendicular to this, indicated by the $(d-1)$-dimensional vector r. Therefore, we write $x = (t, r)$ for a site on the lattice. The relative weight of a
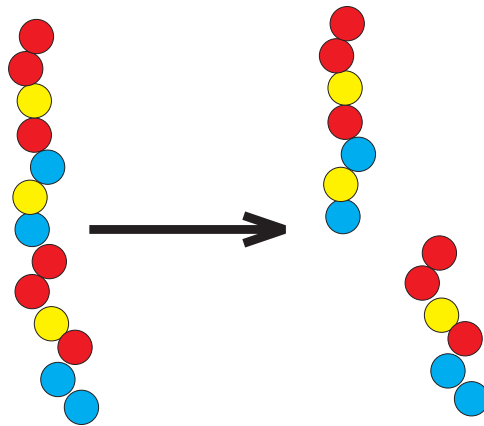
Figure 2. Breaking of a macromolecule.
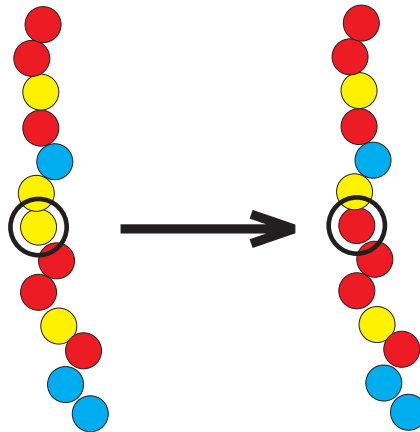
Figure 3. Change of the type of a constituting molecule inside a macromolecule, caused by a collision with a molecule from the surrounding liquid.

nearest-neighbour molecular pair at x, x' of type α and α' is

$$w_{x,x'}^{\alpha\alpha'} = \frac{1}{N} w_{x,x'} = \begin{cases} J/N & \text{if } t' = t+1, \quad r = r' \text{ or } r, r' \text{ nearest neighbours} \\ 0 & \text{otherwise.} \end{cases}$$

The directed chain-like structure is given by the matrix w that is asymmetric in the direction t but symmetric with respect to r. x' is assumed to be 'later' along the direction of the MM. w is also symmetric with respect to α, α'. Thus α and r are of the same type in contrast to t.

The statistical distribution of configurations of MM has to take into account a strong repulsive interaction between molecules of the same type at the same site. This can conveniently be described by a commutative algebra of nilpotent variables $\eta_x^{\alpha,\sigma}$ (i.e., $(\eta_x^{\alpha,\sigma})^l = 0$ if $l > 1$) with $\alpha = 1, 2, \ldots, N$ and $\sigma = 1, 2$ [5]. Each molecule, characterized by its position $x = (r, t)$ on the lattice and by its chemical character (α, σ), is described by a variable. Since $\{\eta_x^{\alpha,\sigma}\}$ are nilpotent, the repulsive interaction results in an exclusion principle for molecules of the same type. The atoms are considered as classical objects. Consequently, products of $\eta_x^{\alpha,\sigma}$ must be

commutative. A segment of the MM (i.e. a bond state of two molecules or a dimer) reads

$$w_{x,x'}^{\alpha\alpha'}\eta_x^{\alpha,1}\eta_{x'}^{\alpha',2}.$$

A linear mapping of the algebra $\{\eta_x^{\alpha,\sigma}\}$ to the complex numbers can be written as an integral on a lattice Λ with [6]

$$\int \prod_{x\in\Lambda'\subseteq\Lambda}\prod_{\alpha\in i_x}\prod_{\sigma\in j_{x,\alpha}}\eta_x^{\alpha,\sigma} = \begin{cases} 1 & \text{if } \Lambda' = \Lambda, \quad i_x = 1,2,\ldots,N, \quad j_{x,\alpha} = 1,2 \\ 0 & \text{otherwise.} \end{cases}$$

The integral vanishes if the product is incomplete with respect to the lattice Λ, to $\alpha = 1,2,\ldots,N$ or to $\sigma = 1,2$. It can be used to write the statistical weight P_I of a configuration I of MM as

$$P_I = \frac{1}{Z}\int W_I$$

with

$$W_I = \prod_{(x,\alpha;x',\alpha')\in I}\frac{1}{N}w_{x,x'}\eta_x^{\alpha,1}\eta_{x'}^{\alpha',2}\prod_{(x,\alpha)}(1+\mu\eta_x^{\alpha,1}\eta_x^{\alpha,2}),$$

where $(x,\alpha;x',\alpha')$ is a bond between two CM of type α and α' at sites x and x'. The normalization Z is the sum over all possible configurations $\{I\}$:

$$Z = \sum_{\{I\}}\int W_I.$$

Finally, we impose periodic boundary conditions on the lattice. This implies a system on a torus with closed loops of MM. However, the choice of boundary conditions should not be crucial for the local properties of the MM like the density of the MM, $G_{xx'}$, $H_{xx'}$ or $S_{\alpha\alpha'}$.

3. Functional-integral representation

Introducing analytic functions of the nilpotent field and using the properties of the integral, we may write for the normalization

$$Z = \sum_{\{I\}}\int W_I = \int \exp\left[\sum_{x,x'}\sum_{\alpha,\alpha'=1}^{N}\left(\frac{1}{N}w_{x,x'}+\mu\delta_{x,x'}\delta_{\alpha,\alpha'}\right)\eta_x^{\alpha,1}\eta_{x'}^{\alpha',2}\right]. \tag{1}$$

In other words, we can replace under the integral

$$\sum_{\{I\}}W_I \to W = \exp\left[\sum_{x,x'}\sum_{\alpha,\alpha'=1}^{N}\left(\frac{1}{N}w_{x,x'}+\mu\delta_{x,x'}\delta_{\alpha,\alpha'}\right)\eta_x^{\alpha,1}\eta_{x'}^{\alpha',2}\right]$$

in Z. Another important consequence of the properties of the nilpotent field is that we can calculate the probability that a point x,α belongs to a MM configuration I:

$$\frac{1}{Z}\int(1-\mu\eta_x^{\alpha,1}\eta_x^{\alpha,2})W_I = \begin{cases} P_I & \text{if } (x,\alpha)\in I \\ 0 & \text{otherwise.} \end{cases}$$

After summation over all configurations $\{I\}$, the probability of a molecule of type α at site x being part of any MM is

$$\sum_{\{I\}}P_I = \sum_{\{I\}}\frac{1}{Z}\int(1-\mu\eta_x^{\alpha,1}\eta_x^{\alpha,2})W_I = \frac{1}{Z}\int(1-\mu\eta_x^{\alpha,1}\eta_x^{\alpha,2})W.$$

This is also the local density $n(\mu)$ of an MM. In a similar manner, we can express the correlations in terms of normalized integrals

$$\langle \cdots \rangle = \frac{1}{Z} \int \cdots W$$

as correlations in real space

$$G_{xx'} = \langle \eta_x^{\alpha,1} \eta_{x'}^{\alpha,2} \rangle = -C(x\alpha; x'\alpha),$$
$$H_{xx'} = \langle \eta_x^{\alpha,2} \eta_{x'}^{\alpha,1} \rangle = -C(x'\alpha; x\alpha),$$

and correlations in α space

$$S_{\alpha\alpha'} = \langle \eta_x^{\alpha,1} \eta_x^{\alpha',2} \rangle = -C(x\alpha; x\alpha').$$

These are correlation functions in real space and in α space, respectively. The density of MM can also be expressed as

$$n(\mu) = 1 - \mu C(x\alpha; x\alpha).$$

A lengthy but straightforward calculation [5] shows that the normalization Z in equation (1) can be written as an functional integral with respect to complex fields φ and χ as

$$Z = \int e^{-NS} \prod_x d\varphi_x \, d\chi_x. \tag{2}$$

with

$$S = (\varphi, (1+w)^{-1}\bar{\varphi}) + (\chi, \bar{\chi}) - \sum_x \log[\mu + (\varphi_x + i\chi_x)(\bar{\varphi}_x + i\bar{\chi}_x)].$$

An analogous calculation for the correlation functions leads to

$$\langle \cdots \rangle = \frac{1}{Z} \int \cdots e^{-NS} \prod_x d\varphi_x \, d\chi_x. \tag{3}$$

The correlation functions then become

$$C(x\alpha; x\alpha) = \left\langle \left[\mu + (\varphi_x + i\chi_x)(\varphi_x^* + i\chi_x^*) \right]^{-1} \right\rangle$$

$$C(x\alpha; x'\beta) = \mu \left\langle \frac{(\varphi_x + i\chi_x)(\varphi_{x'}^* + i\chi_{x'}^*)}{[\mu + (\varphi_x + i\chi_x)(\varphi_x^* + i\chi_x^*)][\mu + (\varphi_{x'} + i\chi_{x'})(\varphi_{x'}^* + i\chi_{x'}^*)]} \right\rangle$$

for $(x, \alpha) = (x', \beta)$.

3.1. $1/N$ expansion

The integrals in equations (2) and (3) depend on the number of different molecule types N only through the prefactor N in the exponential. This enables us to perform a saddle-point integration for large values of N [5]. The result is a $1/N$ expansion. The homogeneous saddle point of the integral in equation (2) satisfies the relation

$$i\chi = -\tfrac{1}{2}\varphi, \qquad i\bar{\chi} = -\tfrac{1}{2}\bar{\varphi}.$$

Then there is a trivial solution $\varphi_0 = \bar{\varphi}_0 = \chi_0 = \bar{\chi}_0 = 0$ and a non-trivial solution with

$$\varphi_1 \bar{\varphi}_1 = 4(1 - \mu).$$

An expansion of the exponent in Z in terms of the deviations around the saddle point

$$\phi_x = \varphi + i\chi + \delta\phi_x \qquad \text{with } \delta\phi_x := \delta\varphi_x + i\delta\chi_x, \qquad \delta\phi_x^* := \delta\varphi_x^* + i\delta\chi_x^*$$

provides corrections in powers of $1/N$. For a non-zero density of MM (i.e. $\zeta \equiv \mu^{-1} > 1$), an expansion up to second order leads to

$$
\begin{aligned}
C(x\alpha; x'\beta) \approx \zeta^{-1} - \zeta^{-2} + \frac{1}{\zeta^3}[&-(\zeta-1)\langle\delta\phi_x^2\rangle - 2(\zeta-1)\langle\delta\phi_x\delta\phi_x^*\rangle \\
&- (\zeta-1)\langle\delta\phi_x\delta\phi_{x'}\rangle + \langle\delta\phi_x\delta\phi_{x'}^*\rangle + \sqrt{\zeta(\zeta-1)}\langle\delta\phi_x\rangle + (\zeta-1)^2\langle(\delta\phi_x^*)^2\rangle \\
&+ (\zeta-1)^2\langle\delta\phi_x^*\delta\phi_{x'}\rangle - (\zeta-1)\langle\delta\phi_x^*\delta\phi_{x'}^*\rangle \\
&- (\zeta-1)\sqrt{\zeta(\zeta-1)}\langle\delta\phi_x\rangle + (\zeta-1)^2\langle\delta\phi_{x'}^2\rangle - 2(\zeta-1)\langle\delta\phi_{x'}\delta\phi_{x'}^*\rangle \\
&- (\zeta-1)\sqrt{\zeta(\zeta-1)}\langle\delta\phi_{x'}\rangle - (\zeta-1)\langle(\delta\phi_{x'}^*)^2\rangle + \sqrt{\zeta(\zeta-1)}\langle\delta\phi_{x'}^*\rangle].
\end{aligned}
$$

The first term is the large-N limit; all other terms either vanish because of $\langle\delta\phi_x\rangle = 0$ or are of order $1/N$ because of $\langle\delta\phi_{x'}^2\rangle = \mathrm{O}(1/N)$. The saddle-point calculation leads to the result that the first-order terms vanish, and the second-order terms can be expressed as the integrals

$$
\begin{aligned}
\Phi(r-r', t-t') &:= \langle\delta\phi_x\delta\phi_{x'}\rangle = \langle\delta\phi_x^*\delta\phi_{x'}^*\rangle \\
&\approx \frac{1}{N} \int \int_0^{2\pi} \frac{B(k)^2(\zeta-1)}{2B(k)\cos(\omega) - \zeta + (\zeta-2)B(k)^2} \, \mathrm{e}^{\mathrm{i}[k(r-r')-\omega(t-t')]} \frac{\mathrm{d}\omega}{2\pi} \frac{\mathrm{d}^{d-1}k}{(2\pi)^{d-1}}
\end{aligned}
$$

(4)

$$
\begin{aligned}
\hat{\Phi}(r-r', t-t') &:= \langle\delta\phi_x\delta\phi_{x'}^*\rangle \\
&\approx \frac{1}{N} \int \int_0^{2\pi} \frac{B(k)(B(k) - \zeta\cos(\omega))}{2B(k)\cos(\omega) - \zeta + (\zeta-2)B(k)^2} \mathrm{e}^{\mathrm{i}[k(r-r')-\omega(t-t')]} \frac{\mathrm{d}\omega}{2\pi} \frac{\mathrm{d}^{d-1}k}{(2\pi)^{d-1}}
\end{aligned}
$$

(5)

with the function $B(k) := 1 - J + \frac{J}{d-1}\sum_{j=1}^{d-1}\cos k_j$.

4. Discussion of the results

The density of MM is

$$
n(\mu) = \begin{cases} 1 - \mu + \mathrm{O}(1/N) & \text{for } \mu < 1 \\ 0 & \text{for } \mu > 1. \end{cases}
$$

It increases with decreasing N, as shown in figure 4. This means that it is more likely to form a MM if the number of available kinds of molecules is reduced.
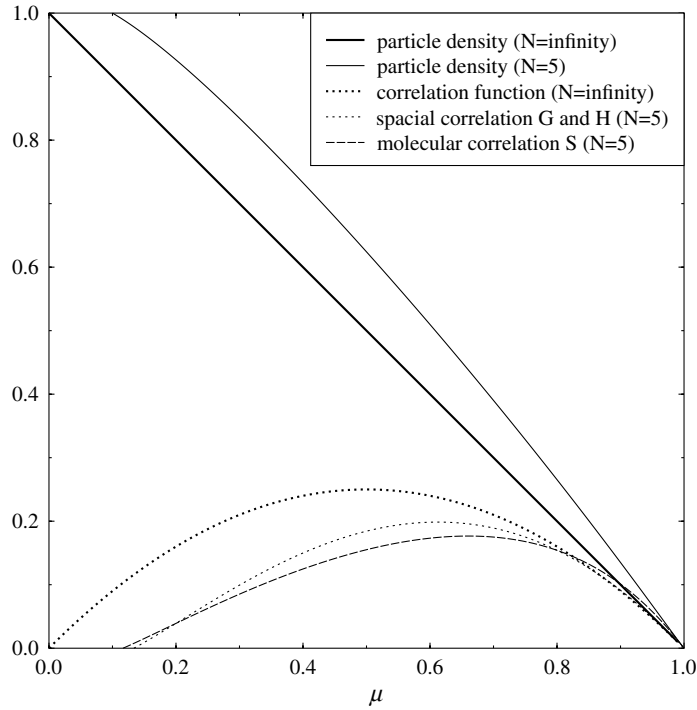
How likely it is to break a MM depends on the position of the remaining ends $x - x'$. The corresponding quantity $G_{xx'}$ decays with $|x - x'|$ which indicates that the molecules have a tendency to form a MM. However, it does not vanish asymptotically but stays nonzero for $\zeta > 1$ and $t' = t$:

$$
G = \lim_{|r-r'|\to\infty} C(x\alpha; x'\alpha) \approx (1 - \zeta^{-1})\left(\zeta^{-1} + \left[-\frac{4}{\zeta^2}\hat{\Phi}(0,0) + \frac{2}{\zeta}\left(1 - \frac{2}{\zeta}\right)\Phi(0,0)\right]\right).
$$

The ζ dependence of this asymptotic result is shown in figure 4. The same property is valid for the creation of a finite chain of molecules with end points x, x', described by $H_{xx'}$.

The correlation in α space $S_{\alpha\alpha'}$, describing how likely it is to change the type of molecules in the MM, does not depend on the specific pair α, α' as long as $\alpha = \alpha'$ because the model connects all $\alpha = \alpha'$ with equal probability:

$$
S = C(x\alpha; x\alpha') \approx \zeta^{-1} - \zeta^{-2} + \frac{1}{\zeta}\left(1 - \frac{6}{\zeta} + \frac{6}{\zeta^2}\right)\hat{\Phi}(0,0) + \frac{2}{\zeta}\left(1 - \frac{1}{\zeta}\right)^2\Phi(0,0).
$$

Figure 4. The density of macromolecules and the correlations for breaking a macromolecule (G) and replacing a molecule in a macromolecule (S).

The $N \to \infty$ terms of both correlations G and S agree but there is a difference in terms of order $1/N$ (see figure 4). It is unlikely to break an MM and separate the pieces at low and at high density. This is a consequence of the fact that there are either too few MM or too many at high densities such that the separation of the pieces is difficult. It is remarkable that already at densities less than 1 the separation is completely blocked, as indicated by the vanishing G for $n_c \leqslant n \leqslant 1$. Moreover, the replacement of a CM is most likely at intermediate densities but suppressed at low and high densities. Again, there is a critical density $n'_c < 1$ such that the substitution is blocked for $n'_c \leqslant n \leqslant 1$. This can be understood as an effect of interaction: since there is a exclusion between molecules of the same type, we can only substitute if there is no contribution of this molecule from another MM. At high densities it is very unlikely that a site is not occupied yet by a CM of a specific type from another MM.

4.1. Algebraic decay of the correlation function

To analyse asymptotic behaviour of the correlation function for large distances $|r - r'|$ of a three-dimensional system, we first have to perform the ω-integration in equations (4) and (5). This leads to

$$\Phi(r - r') = \frac{1}{N} \int \frac{-B(k)(\zeta - 1)}{\sqrt{((\zeta - 2)B(k)^2 - \zeta)^2 - 4B(k)^2}} e^{ik(r-r')} \frac{\mathrm{d}^2 k}{(2\pi)^2},$$

$$\hat{\Phi}(r - r') = \frac{1}{N} \int \left[-\frac{\zeta}{2} + \frac{\frac{\zeta^2}{2} + B(k)^2\left(-1 + \zeta - \frac{\zeta^2}{2}\right)}{\sqrt{((\zeta - 2)B(k)^2 - \zeta)^2 - 4B(k)^2}} \right] e^{ik(r-r')} \frac{\mathrm{d}^2 k}{(2\pi)^2}.$$

The spatially decaying part of the correlation function is given as

$$\frac{2}{\zeta^2}\left(1-\frac{1}{\zeta}\right)\Phi(r-r')+\frac{1}{\zeta}\left(1-\frac{2}{\zeta}+\frac{2}{\zeta^2}\right)\hat{\Phi}(r-r').$$

For small values of k we can assume $B(k)\approx 1$ in the numerators of $\Phi(r-r')$ and $\hat{\Phi}(r-r')$. The square root in the denominators can be approximated by using $B(k)=1-\frac{k^2}{2}+O(k^4)$:

$$\sqrt{((\zeta-2)B(k)^2-\zeta)^2-4B(k)^2}=\sqrt{4(\zeta-1)k^2+O(k^4)}\approx 2\sqrt{\zeta-1}|k|.$$

Therefore, the asymptotic behaviour of the correlation function is proportional to

$$\text{const}\int\frac{e^{ik\cdot(r-r')}}{|k|}\frac{d^2k}{(2\pi)^2}\propto \text{const}\int_0^{2\pi}\int_0^\infty e^{ik|r-r'|\cos\phi}\,dk\,d\phi$$

$$=|r-r'|^{-1}\underbrace{\text{const}\int_0^{2\pi}\int_0^\infty e^{i\kappa\cos\phi}\,d\kappa\,d\phi}_{=\,\text{const}}\propto |r-r'|^{-1},$$

where k has been replaced by $\kappa=k|r-r'|$. Thus the spatial correlations decay like $1/|r-r'|$ for $d=3$.

5. Conclusions

A statistical ensemble of directed MM, constructed with a random sequence of N different molecules, has been treated for large values of N in terms of a $1/N$ expansion. We have studied the robustness of the MM with respect to breaking and replacement of individual CM. Breaking a MM and taking the pieces apart in space reveals a long-range correlation with an algebraic decay with the inverse distance. The properties in space and in α are identical in the $N\to\infty$ limit but different in $1/N$.

At high densities the broken MM cannot be separated to arbitrarily large distances due to the interaction between different MM. The replacement of individual CM is most likely at intermediate densities but strongly suppressed at high densities. This is also a consequence of the repulsive interaction between the molecules of the same type.

References

[1] de Gennes P G 1979 *Scaling Concepts in Polymer Physics* (Ithaca, NY: Cornell University Press)
[2] Li Y-Q *et al* 2004 *Preprint* q-bio.BM/0408024
[3] Banavar J R and Maritan A 2003 *Rev. Mod. Phys.* **75** 23
[4] Banavar J R 2004 *Preprint* q-bio.BM/0410031
[5] Ziegler K 1991 *Physica* A **179** 301
[6] Berezin F A 1966 *Method of Second Quantization* (New York: Academic)