$$A = \begin{bmatrix} 0 & -\kappa & \omega \\ \kappa & 0 & -\tau \\ -\omega & \tau & 0 \end{bmatrix} \tag{2.86}$$

with

$$A_{ij} = \sum_k \epsilon_{ijk}\omega_k \tag{2.87}$$

we can develop also a numerical procedure to generate curves that have a prescribed spontaneous curvature with fluctuations. For this let $v^x = (t_1^x, t_2^x, t_3^x,)$ etc., then we write the equation 2.83 as

$$\frac{dv^i}{ds} = Av^i \tag{2.88}$$

which can be discretized as

$$v^i(s + ds) = Ov^i(s) \tag{2.89}$$

with

$$O = (1 + \frac{ds}{2}A)(1 - \frac{ds}{2}A)^{-1} \quad . \tag{2.90}$$

Note that O is an orthogonal matrix guaranteeing that we consistently generate orthogonal triads.
In figure 2.14 the case for a curve is depicted where the spontaneous curvature is set such that a circle is preferred.

2.1.4 Self-Avoiding Chains

Important Concepts: self avoidance, scaling exponent, Flory argument, lattice chains, n-vector model

The above models all lack the excluded volume interaction between the monomers. Indeed, the ubiquitous polymer model is a self-avoiding random walk.
Let $U(\mathbf{R}_n - \mathbf{R}_m)$ be a monomer-monomer interaction potential assumed to handle the excluded volume. We treat the polymer as gas of disconnected monomers

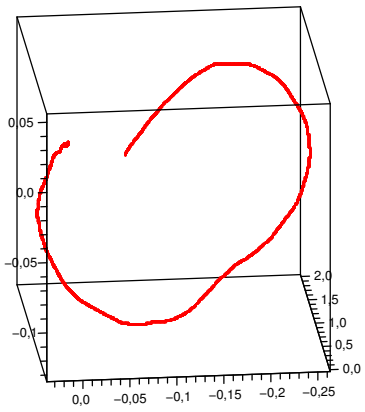

Figure 2.14: Example of a random walk with torsion of length 1 and parameters $\kappa = 2\pi$ and $\tau = 0$

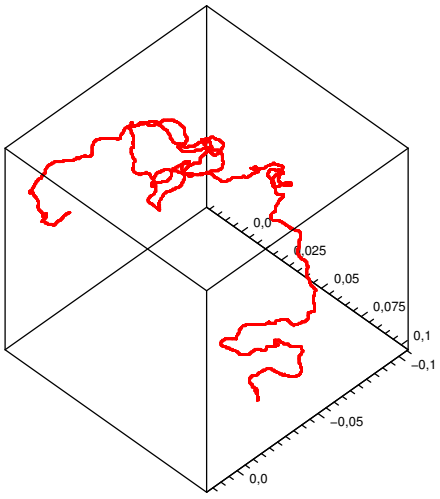

Figure 2.15: Example of a random walk with torsion. The parameters were: 2pi-0-0 ds0001 100.000000.txt

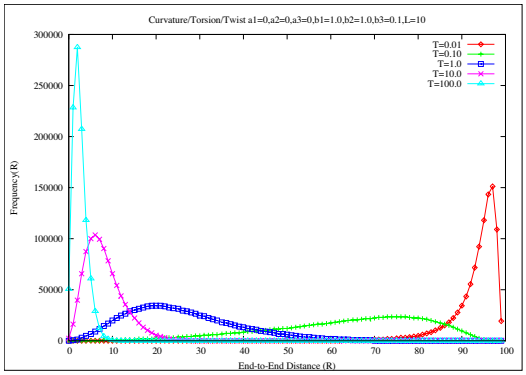

Figure 2.16: Distribution

confined within the same volume V as the polymer coil. If $\bar{R}$ denotes the average extent of the chain then the chain occupies a volume $V = \bar{R}^3$. Let ΔF_c be the term that arises from the chemical work done by initially "preparing" the monomers in the box. This must be independent of $\bar{R}$. Further the free energy term dependent on $\bar{R}$ must be extensive and thus dependent on V. Since this term involves an interaction between monomers it must depend on the concentration squared

$$\Delta F_R = v k_B T c^2 V \quad . \tag{2.91}$$

The free energy cost of putting N excluded monomers into a system with volume V, at average concentration $c = N/V$, is then given by

$$\Delta F_{\text{ex}} = \Delta F_c + \Delta F_R \quad . \tag{2.92}$$

The total free energy change includes the entropic term (2.27) and the term coming from the excluded volume (neglecting terms not dependent on $\mathbf{R}$)

$$\Delta F = \Delta F_e + \Delta F_R \approx k_B T \left(\frac{\bar{R}^2}{N b^2} + v \frac{N^2}{\bar{R}^3} \right) \quad . \tag{2.93}$$

This must be minimized with respect to $\bar{R}$, i.e., we have to solve

$$\frac{\partial \Delta F}{\partial \bar{R}} = 0 \tag{2.94}$$

and find

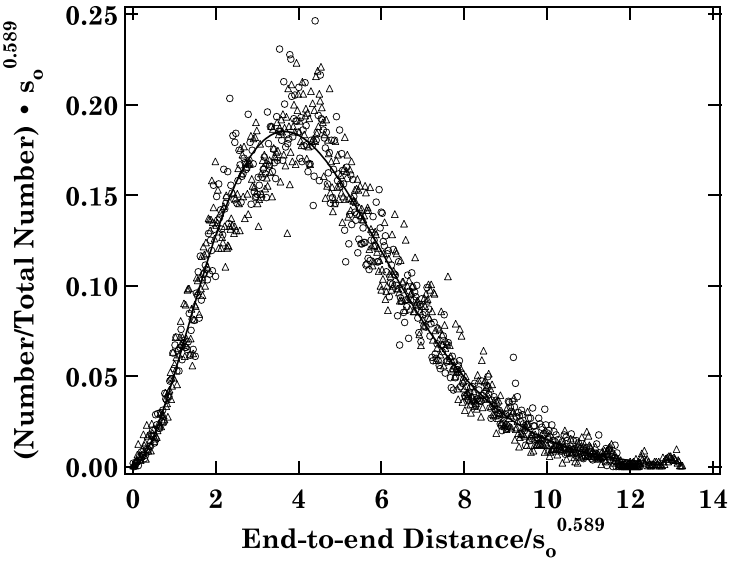

Figure 2.17: End-to-end distance

$$\bar{R} \approx \left(\frac{v}{b^3}\right)^{1/5} bN^{3/5} \quad . \tag{2.95}$$

If we write

$$< R > \propto N^{\nu} \tag{2.96}$$

then we obtain for the gaussian case (without self avoidance) $\nu = 1/2$ and with excluded volume (within the above mean field theory) $\nu = 3/5$.

Further Reading 2.1.6 (Lattice Chains)

The self-avoiding random walk (SAW) on a periodic lattice was considered by Orr [38] as a model of a polymer chain. Such a self-avoiding random walk is shown in figure 2.19. In one dimension the problem is trivial and unsolved in higher dimensions.

Let c_N denote the number $n - step$ self-avoiding walks (SAW) (equivalent upon translation!). We can easily enumerate on the square lattice $c_1 = 4$, $c_2 = 12$, $c_3 = 36$ and $c_4 = 100$. In general it is believed to be [39, 40, 41]

$$c_N \approx A\mu^N N^{\gamma-1} \tag{2.97}$$

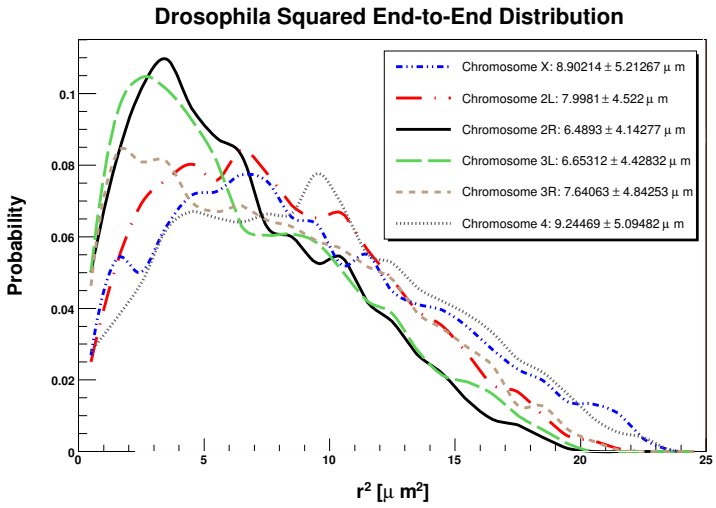

Figure 2.18: Histogram representing the distribution of the end-to-end distance for two different contour lengths (548nm circles, 748nm triangles) and how they collapse onto each other. Taken from Dietler et. al. PRL 95, 158105 (2005)

with γ being a universal exponent ($d = 2\ \gamma = 32/43,\ d = 3\ \gamma \approx 7/6,\ d \geq 4\ \gamma = 1$) and μ the *connectivity constant* giving the average number of available steps for an infinitely long walk.
For the partition function we have

$$Z_N \sim q_\mu^N N^{\gamma-1} \quad q_{\text{eff}} < q(\Lambda) \tag{2.98}$$

and thus for the average end-to-end distance

$$< R_e^2 > \propto N^{2\nu} \tag{2.99}$$

with $\nu \approx 0.59$ (in 3d) and $\gamma \approx 1.158$ (in $3d$) from numerical calculations.

Further Reading 2.1.7 (Self-Avoiding Walks and the n-Vector Model)
We will now work out the connection between spin-models and the self-avoiding random walk. For this we consider a lattice Λ with N sites. Each lattice site is occupied by an n-component spin of length $\sqrt{n}$

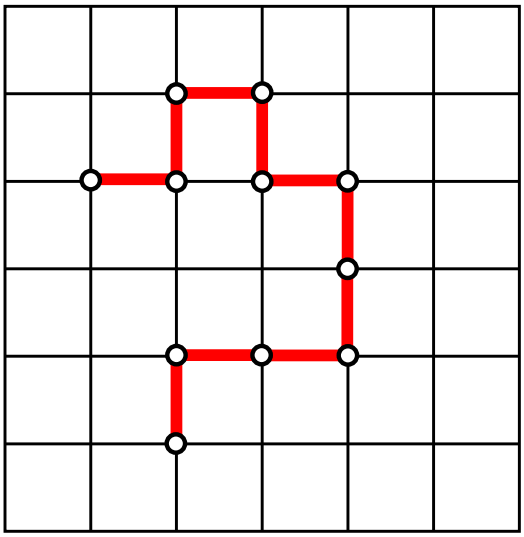

Figure 2.19: A self-avoiding random walk (SAW)

$$\mathbf{s}_i = (s_{i1}, s_{i2}, ..., s_{in}) \quad , \mathbf{s}_i^2 = n \tag{2.100}$$

with a Hamiltonian

$$H = -\frac{J}{T} \sum_{<i,j>} \mathbf{s}_i \mathbf{s}_j - \frac{h}{T} \sum_i \mathbf{s}_{i1} \quad , \tag{2.101}$$

where $J > 0$, and $<>$ denotes summing of nearest neighbour pairs. For $n = 1$ this yields the Ising model and for $n = 3$ we recover the classical Heisenberg model (see section 2.1.3)
First we look at the moments of an isolated n-vector spin $\mathbf{s}$ ($K = J/T = 0, h/T = 0$) which varies on an n-dimensional sphere with radius $\sqrt{n}$. Let us denote the normed integration over the sphere by $<>_0$

$$\frac{\int d^n s \delta(\mathbf{s}^2 - n) f(\mathbf{s})}{\int d^n s \delta(\mathbf{s}^2 - n)} \equiv < f(\mathbf{s}) >_0 \tag{2.102}$$

from which we obtain

$$n = < \sum_{\alpha=1}^{n} s_\alpha^2 >_0 = n < s_1^2 >_0 \quad , \quad < s_\alpha s_\beta >= \delta_{\alpha\beta} \tag{2.103}$$

independent of n. Further we find

$$< f(\mathbf{s}_\alpha)g(\mathbf{s}_\beta) >_0 = < f(\mathbf{s}_\alpha) >_0 < g(\mathbf{s}_\beta) >_0 \quad \text{for} \quad \alpha \neq \beta \quad . \tag{2.104}$$

For the fourth moment we find

$$< s_1^4 >_0 = 3 < s_1^2 s_2^2 >_0 \quad . \tag{2.105}$$

Thus

$$\begin{aligned} n^2 &= < \left(\sum_{\alpha=1}^{n} s_\alpha^2 \right)^2 >_0 & (2.106)\\ &= n < s_1^4 >_0 + (n^2 - n) < s_1^2 s_2^2 >_0 & (2.107)\\ &= (n^2 + 2n) < s_1^2 s_2^2 >_0 & (2.108) \end{aligned}$$

and

$$< s_1^2 s_2^2 >_0 = \frac{n}{2+n} \quad , < s_1^4 >_0 = \frac{3n}{2+n} \quad . \tag{2.109}$$

To make the connection to self-avoiding random walks we shall take the limit $n \to 0$ for the case $h = 0$ and $T > T_c$, i.e. temperatures above the critical point and show that the spin-correlation-function corresponds to the probability distribution of the end-to-end distance of the SAW.
For $n = 1, 2, ...$ we have for the spin-correlation

$$\Gamma(r) = < s_{01} s_{r1} > = < s_{01} s_{r1} e^{\bar{H}} >_0 \; / \; < e^{\bar{H}} >_0 \quad , \tag{2.110}$$

which can be written as the sum of products of single spin components due to the expansion

$$e^{\bar{H}} = \prod_{<i,j>} e^{J/T \mathbf{s}_i \mathbf{s}_j} = \prod_{<i,j>} (1 + \frac{J}{T} \mathbf{s}_i \mathbf{s}_j + ...) \quad . \tag{2.111}$$

Since emanating from each lattice point we have only 0 or 2 spin-factors in $<>_0$ for $n \to 0$ we find that the numerator has only contributions of the form shown in figure 2.20.
Thus

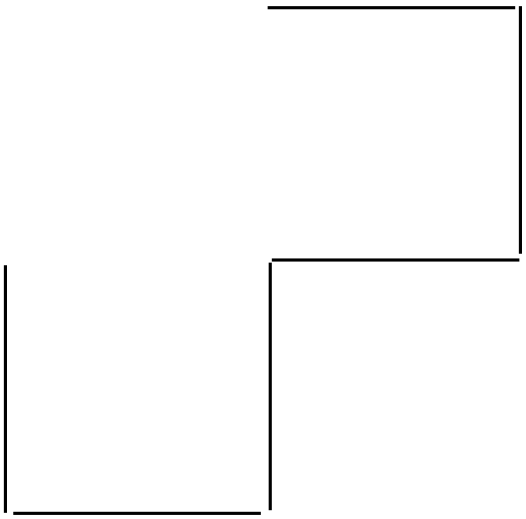

Figure 2.20: self-avoiding chains contributing to the n-vector model in the limit $n \to 0$.

$$\Gamma_{n\to 0}(r) = \sum_{N=0}^{\infty} \zeta_N(r) \quad , \tag{2.112}$$

where $\zeta_N(r)$ is the number of self-avoiding random walks with N steps starting at the origin and ending at r.

Further Reading 2.1.8 (Modeling of gene silencing)

One case of gene silencing that shares common features with heterochromatin silencing involves the proteins of the Polycomb group (PcG) [56]. PcG proteins are highly conserved regulatory factors that are responsible for the maintenance of the silent state of important developmental genes, such as homeotic genes. In Drosophila melanogaster, PcG proteins form multimeric complexes and regulate their genes through binding to chromosomal regulatory elements named PcG response elements (PREs). This silencing involves repressive modifications on the target chromatin. In addition, it has been observed that silencing via PcG proteins and PREs is enhanced by the presence of multiple copies of PRE-containing elements in the nucleus. These copies may, but do not have to be on the same chromosome. Long-distance pairing between these two loci, which brings them closer together than they would usually be, leads to strong repression of the genes

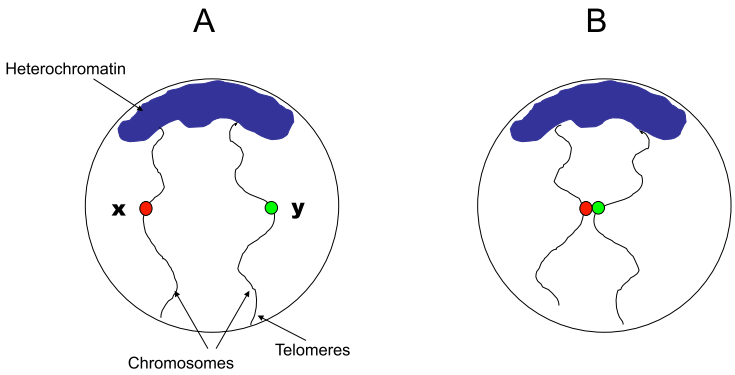

Figure 2.21: Schematic representation of long-distance pairing between two loci X and Y in interphase nuclei. Nuclei are represented in a Rabl configuration, in which centromeres are assembled near the apical pole of the nucleus, whereas the telomeres point toward the basal pole. The wavy lines represent the chromosomes projecting from centromeric heterochromatin toward the basal pole of the nucleus. Colored dots represent independent loci that are distant in a normal situation (A) but can come in close proximity upon integration of PRE-containing elements (B). This phenomenon leads to enhanced silencing and is dependent on PcG proteins.

they control (Bantignies et al [56]). This type of regulation represents silencing by geometrical closeness, established in interphase nuclei (see Fig. 2.21).

In this report, we tried to model long-distance interactions among PREs, with the long term goal to build predictive models for proximity and interactions of chromosomal domains. Within the model, chromosomes were assigned a Rabl configuration [57] in the nucleus, a situation which is present in Drosophila embryonic nuclei. We calculated the expected distance distribution of the two loci in question and compared this distribution to experimental results that were obtained previously.

To model this situation the two arms of the chromosomes carrying the gene loci are represented by two polymer chains. These chains are built up of ellipsoidal monomer segments with a ratio of $\frac{1}{20}$ between the half-axis (for a detailed description of the physical model used see [58]). The chains have to be non-ideal, because the chromosomes cannot penetrate one another. Hence we have to de-

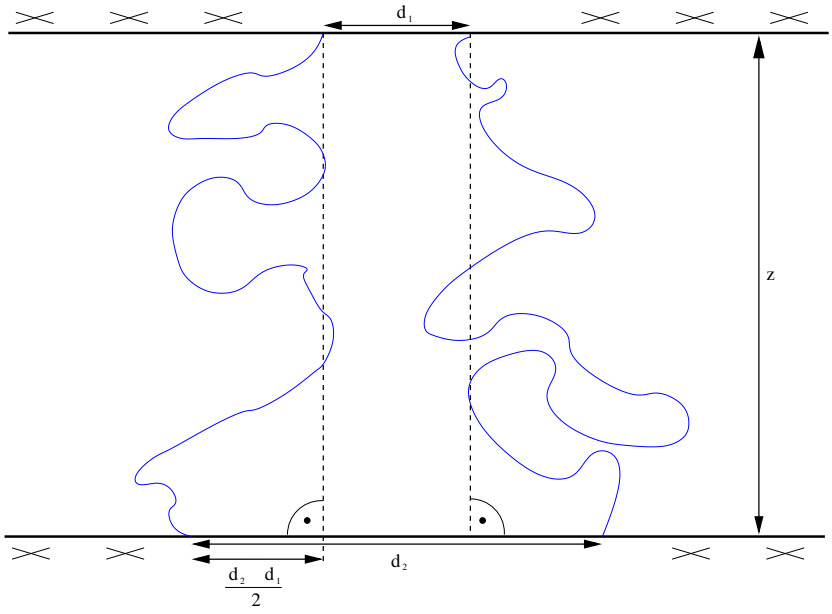

Figure 2.22: Definition of the geometric simulation parameters. This sketch illustrates the meaning of the simulation parameters given in table **??**. The chain contour length is $L = N \times a$.

mand that neither of the two monomers can occupy the same space at the same time. To ensure that our model is comparable to a Rabl configuration, where the DNA is "stretched" from the apical pole of the cell nucleus to its bottom pole (see Fig. 2.21), we need a force acting on polymers which hinders them from curling. The easiest way to do so is to fix the chains between two walls. The distance between these walls should be large enough to ensure the stretching, yet small enough to allow for movement of the chains. In our model the ratio between chain length and wall-to-wall-distance ranges from 1.5 to 5. Both chains have the same length, with the end-points fixed in a plane. The upper and lower end-points are symmetrical (see Fig. 2.22).

Our simulations will be compared with two different sets of experiments. In both cases the two genetic loci in question were positioned on different chromosomes. This makes our model applicable as there is no direct connection between the chromosomal arms we consider. If they were positioned on different arms of the same chromosome, the centromer region might serve as a transmitter, any

”tugging motion” of one locus might be perceptible for the other one. Fixing both chains to different fixpoints on the wall does not incorporate this feature. In both experiments, copies of an element called *Fab-7* (*Fab-7* is an element containing a PRE that regulates an homeotic gene of Drosophila melanogaster) have been inserted into different loci X and Y:

Experiment I used *Drosophila* embryo cells, for which *Fab-7* had been inserted at locus X1 (“*sd*”) and Y1 (“BX-C”). These loci are situated approximately at one third from the top of the nucleus according to F. Bantignies.

Experiment II used *Drosophila* embryo cells, for which *Fab-7* had been inserted at locus X2 (“*sd*”) and Y2 (“38F”). Locus X2 (“*sd*”) is still approximately one third from the top of the nucleus, but locus Y2 (“38F”) is much closer to it ($\approx$ one sixth to one eighth) according to F. Bantignies.

To make the distance distributions obtained through the simulations comparable to these experimental results, we have to scale the data. We may assume that the fixpoints on the walls in our simulation correspond to (loose) binding of the chromosomes to the membrane of the nucleus as outlined in Fig. 2.23. We neglected the fix point distance at the top (d_1) as it is quite small compared to z. As we know the diameter D of the sphere, we can use it to find an approximate scaling relation. In fact, this is the highest scaling factor we may assume that allows us to map the simulation setup into the inside of the nucleus. If we were to consider the less likely scenario where the ends of our chromosomes are not fixed to the cell membrane, but to some other place within the nucleus, our scaling factor would have to be chosen smaller, as the distance between the beginning and the end points of the chains, which we hold fixed, will correspond to less than the diameter of the nucleus. Any larger Kuhn segment length would place at least one of the fixpoints outside the nucleus.

The length c in Fig. 2.23 is most suitable for the derivation. In units of D we have:

$$\begin{aligned} \tan\alpha &= \frac{d_2}{2z} \\ c[D] &= D\cos\alpha, \end{aligned} \tag{2.113}$$

whereas in our arbitrary length scale, whose units we shall call $[a]$, we have:

$$c[a] = \sqrt{z[a]^2 + \frac{d_2[a]^2}{4}} \tag{2.114}$$

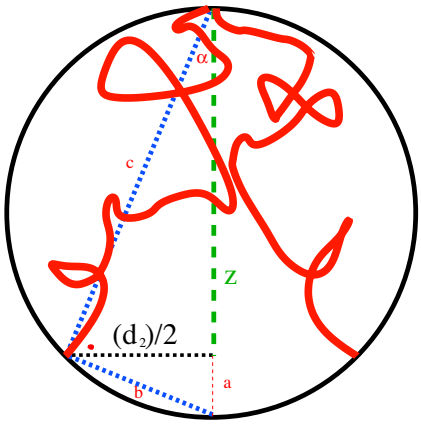

Figure 2.23: Determination of length scale. Here z is the wall-to-wall distance of our simulation setup, d_2 is the distance between the fixpoints on one wall (the distance d_1 on the other wall has been neglected as ≈ 0). The angle α as well as the distances a, b and c have been introduced solely for the sake of calculating the conversion factor from arbitrary length scale to micrometer and have no further meaning.

which, using $D \approx 5\mu m$, leads to

$$1[a] \approx \frac{D\cos(\tan^{-1}(\frac{d_2}{2z}))}{\frac{1}{[a]}\sqrt{z[a]^2 + \frac{d_2[a]^2}{4}}}[\mu m]. \tag{2.115}$$

For the parameter sets A, C and D (see table **??**), where d_2 is much smaller than the wall-to-wall distance z, we may approximate equation (2.115) so that $1[a] \approx \frac{5}{z}[\mu m]$, in case B we have to use the complete formula.
The value of the Kuhn segment length for chromosomes is not yet well known, though most biologists agree that it is of the order of a few hundred nm (see e.g. [**?**]). When we rescale our data according to equation (2.115), we also determine the resultant Kuhn segment and chain contour length (see table 2.2). The Kuhn segment lengths are in the right order of magnitude. To have an immediate comparison to the experimental histograms, we also performed a rebinning of our simulation data into the same ranges Bantignies et al. used in their data processing.
We evaluated the distance distributions between point n_1 on chain 1 and point

Table 2.2: Effective chain and Kuhn segment length for the different parameter sets. The geometry of the system fixed the scaling relations between arbitrary units of the Kuhn segment length to μm.

Number of segments per chain N	Approximate chain length L $[\mu m]$	Resultant effective Kuhn segment length l_k $[nm]$
Parameter Set **A**		
60	10	170
80	13	170
100	17	170
150	25	170
Parameter Set **B**		
60	8	130
80	11	130
100	13	130
150	20	130
Parameter Set **C**		
60	6	110
80	9	110
100	11	110
150	16	110
Parameter Set **D**		
60	10	170
80	13	170
100	17	170
150	25	170

n_2 on chain 2. In accordance with the position of the genes along their chromosomes, for experiment I (*sd* and BX-C) we chose $n_1, n_2 \in [0.3N; 0.4N]$, where $n = 0$ would correspond to the fixpoint on the top wall. For experiment II (*sd* and 38F) we decided to consider ranges of $n_1 \in [0.1N; 0.2N]$ and $n_2 \in [0.3N; 0.4N]$. An example of the resulting graphs for parameter set D can be seen in Figs. 2.24 and 2.25. By visual judgement of these graphs we compared the simulation data to the biological experimental data. When checking for possible agreement between the model and the experiment, we realized that there is most probably the possibility of identifying the distance distributions in the groups with two copies of the gene with distributions from our simulation. We realized that congruence between the curves increased as the number of segments per chain increased. We also found that a larger Kuhn segment length is favorable. In these comparisons we always neglected discrepancies in the very first bin (smallest distances) as the biological model predicts a short-range binding force keeping the genes together once they come within range. This force has not yet been incorporated in the simulation model. On the other hand, the graphs also give strong evidence that no parameter choice within our simple model allows for identification between the control group distributions and the simulation, because the full-width half-maximum (FWHM) of the experimental curves is large although the mean value is relatively small. Parameter choices that could produce a comparable FWHM in the simulation would always place the mean value of the distribution at a much larger value.

2.1.5 Conformations and Energy Landscapes

Important Concepts: energy landscape, random heteropolymer, REM model, glas transition temperature

The conformations of a macromolecule give rise to a complicated potential energy surface. The free energy landscape represents the configuration space of energy and entropy available to a macromolecule. Thus local minima or metastable states, basins of attraction and the saddle points separating them appear. To understand and pave the way for the discussion on protein folding we qualitatively describe the situation. Let us divide the degrees of freedom of the polymer into two sets.

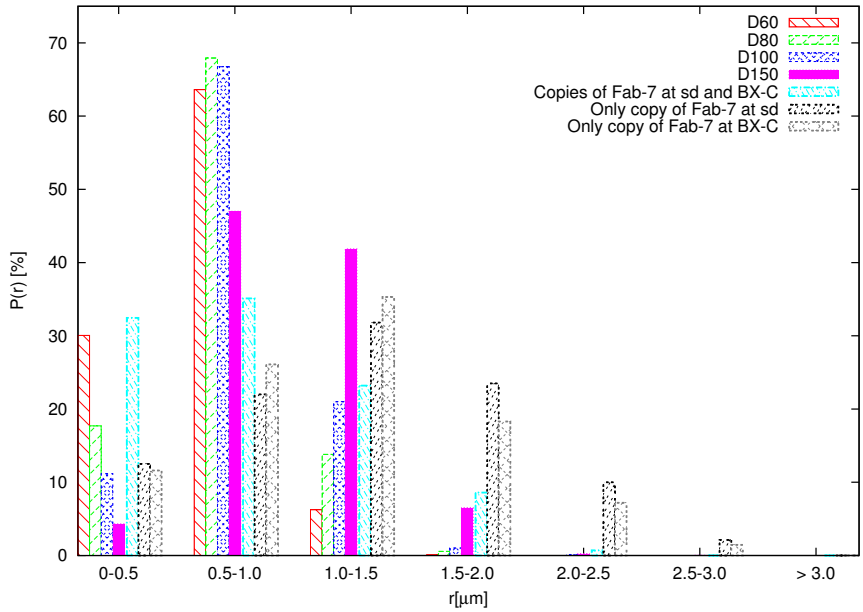

Figure 2.24: Here we see the different distance distributions that involve the *sd* and the BX-C loci. The four bars on the left hand side in each group represent different chain lengths of 60, 80, 100, and 150 Kuhn segment lengths, respectively. In the case shown the Kuhn segment length was 170 nm. The next bar (turquoise) gives the experimental results when two copies of *Fab-7* are present, the black and grey bars stand for the control groups with only one copy.

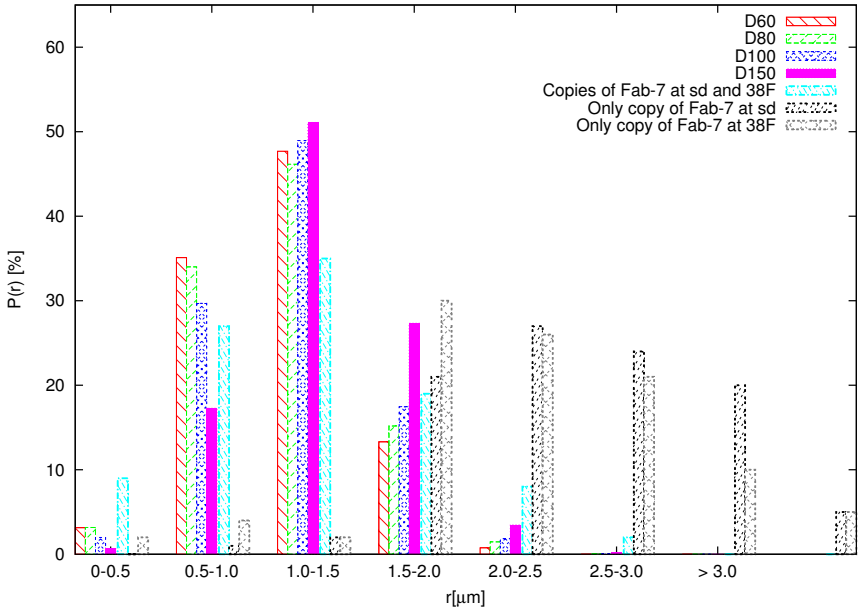

Figure 2.25: In the above graphics we see the different distance distributions that involve the *sd* and 38F loci. The legend is identical to Fig. 2.24.

The first contains the dihedral bond angles $\theta_1, \phi_1, ..., \theta_N, \phi_N$ along the backbone. Into the second set we place everything else from hydrogen bonding, torsion angle energies to rotations of individual side chains. The first set contains the conformational degrees of freedom and the second set the internal degrees of freedom (which allow typical free energy changes on the order of k_BT). The energy landscape represents the $2N$ conformational degrees of freedom. Each configuration is represented by a point on the conformation space such that similar conformations are nearby.

The roughness of an energy landscape may be quantified by the presence of structural hierarchy. Within a closed contour of constant elevation, there exist several closed contours of lower elevation, within each of which are more contours of lower elevation, etc. Landscapes characterized by a hierarchy of sub-valleys within valleys are said to be rugged; trivially hierarchical landscapes, in which each closed contour contains not multiple but a single closed contour of lower elevation, are called smooth.

The conformational energy landscape of a polymer is determined by self-avoidance. To pass to a conformationally near but topologically distant conformation, the polymer must swell and recollapse, overcoming a large energy barrier. A hierarchy of valleys makes the ground state conformation only marginally lower in energy than quasi-degenerate local minima, which act as energetic traps. Further, if the polymer is composed of serval different monomers, as for example in a protein, further constraints arise due to frustration, the inability of chain segments to cooperatively align. Together, these suggest that the conformational landscape of a typical sequence is typically rugged.

We start off with the random energy model (REM) [24, **?**, 25, 26] for the density of states of a heteropolymer. This model aims to describe differences between the energy spectrum of randomly generated sequences, which are unable to fold, and the energy spectrum of the particular set of sequences that fold to a unique, stable, native conformation.

The random heteropolymer (RHP) can be described by the Hamiltonian

$$H_{\mathrm{RHP}}(\{s_i\}, \{\mathbf{R}_i\},) = \sum_{i<j} \epsilon(s_i, s_j)\Delta(|\mathbf{R}_i - \mathbf{R}_j|) \quad . \tag{2.116}$$

Here i counts the monomers (residues) along the chain, $s_i \in \{1, ..., p\}$ is the

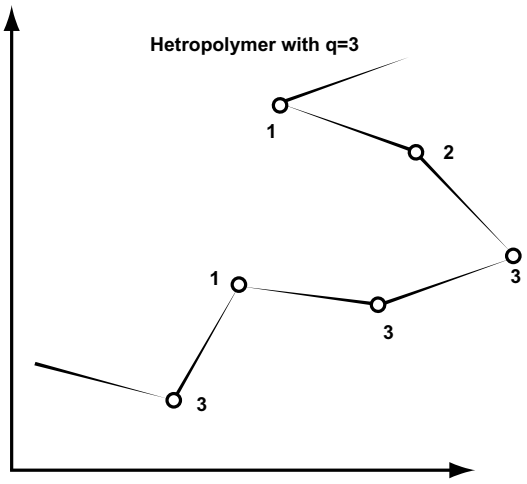

Figure 2.26: Part of a heteropolymer conformation with $q = 3$.

species of monomer and $\mathbf{R}_i$ is the position of monomer i (see figure 2.26). Thus $p = 1$ results in a homopolymer, $p = 2$ into a co-polymer. The special limit $p \to \infty$ stands for a continuous distribution of interaction energies.

It is understood that the connectivity requirement is met.

$\Delta(r)$ includes the excluded volume effects and is assumed to vanish at larger distances.

We now consider a macroscopic system with energy E given by the sum of many microscopic energetic terms ϵ. If individual terms are considered to be statistically independent we can use the central limit theorem to obtain the energy distribution of the system $P(E)$

$$P(E) = \frac{1}{\sqrt{2\pi d}} \exp\left(-\frac{(E - \langle E\rangle)^2}{2d^2}\right) \quad , \tag{2.117}$$

where $d^2 = zNs^2$ with z being the mean number of contacts per monomer and s^2 being the energetic variance of the individual contacts. $< E >= Nz\bar{\epsilon}$ is the average energy. Note that the basic assumption that contact energies are uncorrelated effectively ignores chain connectivity!

The density of microstates is obtained by multiplication of $P(E)$ by the total number of microstates Here with energy given by the sum of individual contacts the density of states is given by

$$n(E) = \gamma^N P(E) = \gamma^N \frac{1}{s\sqrt{2\pi Nz}} \exp\left(-\frac{(E - Nz\bar{\epsilon})^2}{2Nzs^2}\right) \quad , \tag{2.118}$$

where γ is the average number of conformations per monomer taking excluded volume into consideration. This implies that the large majority of conformations will have energy between $Nz\bar{\epsilon} - s\sqrt{Nz}$ and $Nz\bar{\epsilon} + s\sqrt{Nz}$. Further, there exists a critical energy value E_g, obtained from the condition $n(E_g) = 1$

$$E_g = Nz\bar{\epsilon} - Ns\sqrt{2z\ln(\gamma)} = \langle E \rangle - \sigma_E\sqrt{2S_0} \tag{2.119}$$

terms proportional to $\ln N$ were discarded from the above expression as they are small compared to terms proportional to N for large N and $S_0 = N\ln\gamma$ is the conformational entropy of the chain. Below this point the gaps in the spectrum are too large for the chain to change its conformation in the process of thermal fluctuation.

Using eq. 2.118 we have that the entropy is given by

$$S(E) = \ln(n(E)) = N\ln(\gamma) - \ln(b\sqrt{2\pi Nz}) - \frac{(E - Nz\bar{\epsilon})^2}{2Nzs^2} \tag{2.120}$$

and

$$T(E) = \frac{1}{dS/dE} = -\frac{Nzs^2}{E - Nz\bar{\epsilon}} \quad . \tag{2.121}$$

The corresponding critical temperature is given by

$$T_g = T(E_g) = -\frac{Nzs^2}{E_c - Nz\bar{\epsilon}} = s\sqrt{\frac{Nz}{2N\ln(\gamma)}} = \frac{s}{\sqrt{2S_0}} \quad . \tag{2.122}$$

Below T_g the chain will behave like a glass, *freezing* in one of many local minima in its energy surface and never reaching thermal equilibrium. As the energy of the native conformation (the lowest energy in the spectrum), E_N, for a sequence taken at random is expected to be close to E_g, it follows that the native conformation of such a sequence would be thermodynamically stable only at a temperature close to T_g and so would not be able to fold in a reasonable time.

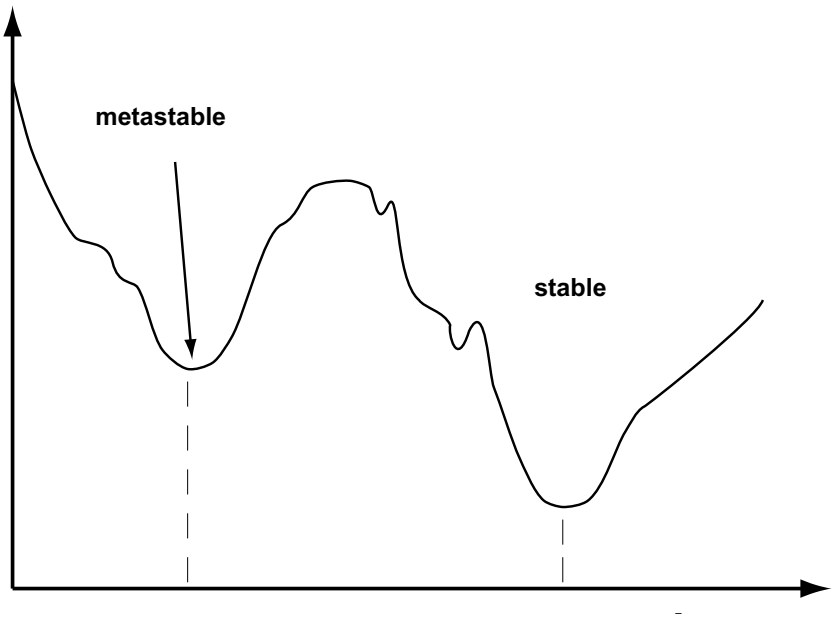

Figure 2.27: Energy landscape

2.1.6 Computational Models

Polymers can be modeled for computational purposes in a variety of ways [27]. Depending on the kind of question and the degree of abstraction, one has the basic choice between a model on a lattice or in continuous space. The bond fluctuation model [28] is one of the prominent representatives for a polymer model on the lattice. The main advantage of this type of models is the computational efficiency due to the restricted configuration space. With increasing computer power it was possible to stay closer to reality by simulating polymers by continuum models. Two widely used models of this class are the bead spring [29] and the united atom model [30].

In both models monomers or parts of them are considered to be represented by spherical force fields. In the united atom model the CH_2 groups are modeled by a spherical force field and the bonded interactions by harmonic forces. In this more atomistic model the anisotropic intermolecular potential functions of polyatomic molecules are constructed using spherical force fields. As an effect the inner degrees of freedom of the molecules like the stiff bonds between the units must also be taken into account. As the Newton equations have to be integrated

such molecular-dynamic simulations are restricted to small time scales.
Other models have been developed in order to adapt an aspherical model to a molecule's geometry i.e. J. Kushick's and B.J. Berne's model [31] and J.G. Gay's and B.J. Berne's model [32]. They consider ellipsoids as a model for molecules and calculate the forces between two interacting ellipsoids as a function of the overlap volume.
The continuous backbone mass model in some sense interpolates between of the united atom model and the bead spring model. On the one hand it tries to stay as close as possible to the chemical realistic structure like the united atom model, but on the other hand it integrates out all the inner degrees of freedom just the same as the bead spring model in order to be computationally efficient. In contrast to these two models it uses *non-spherical* force fields for the non-bonded interaction. The main idea of this approach with a more general form of the force field is to generalize the united atom model in a way that larger atom groups are combined to one construction unit, but the possible anisotropy of these groups is still taken into account. The reasoning is that the topology of the monomer has a strong influence on the physical properties. The simplest anisotropic geometrical object one can think of is an ellipsoid of rotational symmetric form and thus it is considered as the interaction volume of the chemical sequences in our model.
As one wants the force field to degenerate into a sphere with increasing distance, we use a con-focal force field inside this interaction volume:

$$H_{\text{inter}} = V_{\text{abs}} \left(\frac{d_1^{(p)} + d_2^{(p)}}{2} - c \right), \tag{2.123}$$

where $d_1^{(p)}$ and $d_2^{(p)}$ denote the distance of the point $\mathbf{p}$ to the focal points of the ellipsoid and V_{abs} is the absolute potential. In the case of the BPA-PC we take only a repulsive part

$$V_{\text{abs}}(r) = r^{-6} \tag{2.124}$$

into account because from quantum chemical calculations the attractive part proves to be negligible. The calculation of the distances is illustrated in figure 2.28.
The mass of the building units is distributed between the focal points of the ellipsoids, in the hard core region of the con-focal potential. Because of this we perform the crossover from a zero dimensional simulation, i.e. using point masses,

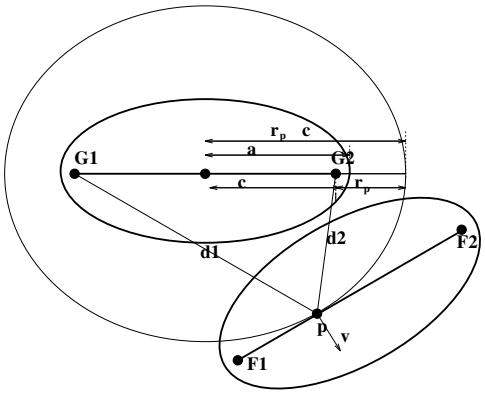

Figure 2.28: Interaction with a con-focal force field

of flexible polymers to a one-dimensional simulation. This has also the effect that no bond-crossing can occur and thus prevents unphysical effects. These one dimensional rods are assumed to be homogeneous. The calculation of the force of such a rod in a con-focal force field leads to an elliptic integral. For this reason we use a Gaussian integration for the computation of this force. In order to form a linear chain the ellipsoids are connected at their focal points.

The parameterization of this model can be generalized to explicitly include side-groups as well as branched polymers, which are linked to the backbone with harmonic potentials in order to have no influence on the moment of inertia of the rod chain.

The main ingredient of the simulation is the mass matrix of our rod-chains. In order to construct it we must calculate first the Lagrangian of a single rod $L_i = T_i - V_i$ with the kinetic energy T_i and the potential energy V_i. The subindex i marks the position of the rods in the chain. This one-dimensional homogeneous rod i has the length l_i starting at $\vec{a}_i$ and ending at $\vec{b}_i$. If we suppose that the rods all have the same mass m and that the velocity of the rod mass scales linearly with the position between the boundaries of the rod, the kinetic energy can be written as

$$T_i = \frac{1}{2}\int_0^{l_i} \frac{m}{l_i}\left(\frac{(l_i - x)\dot{\vec{a}}_i + x\dot{\vec{b}}_i}{l_i}\right)^2 dx = \frac{1}{6}m(\dot{\vec{a}}_i^{\,2} + \dot{\vec{a}}_i\dot{\vec{b}}_i + \dot{\vec{b}}_i^{\,2}). \qquad (2.125)$$

Adding the single terms of the rods building the chain we get the Lagrangian L

of the whole rod chain. The equations of motion of the chain can be calculated from the Lagrange equations of the second kind. Since the equations of motion separate in each direction, we have only to solve three tridiagonal $(N+1)\times(N+1)$ matrices per chain which consist of N rods per time step of the form

$$\mathbf{W}\ddot{\vec{x}} = \vec{F} \tag{2.126}$$

$$\frac{m}{6}\begin{pmatrix} 2 & 1 & 0 & 0 & \dots \\ 1 & 4 & 1 & 0 & \dots \\ 0 & 1 & 4 & 1 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}\begin{pmatrix} \ddot{x}_0 \\ \ddot{x}_1 \\ \ddot{x}_2 \\ \vdots \end{pmatrix} = \begin{pmatrix} F_{10} \\ F_{11}+F_{21} \\ F_{22}+F_{32} \\ \vdots \end{pmatrix} \tag{2.127}$$

with the force F_{ij} on the coordinate j of the flexible point i of the chain

$$F_{ij} = -\frac{\partial V_i}{\partial j} \tag{2.128}$$

and $\ddot{x}_i$ denote the accelerations of the flexible points of the chain. The flexible points are the link points of the ellipsoids and the end points of the rod chain. The sub-indices mark the positions in the chain: 0 and $N+1$ are the end points of the chain and the numbers between them denote the linking points of rods in the chain.

The bonded interactions between neighboring units are given by harmonic length and angle potentials:

$$H_{bond} = \frac{1}{2}k(r-r_0)^2 \tag{2.129}$$

$$H_{angle} = \frac{1}{2}k_\theta(\cos\theta - \cos\theta_0)^2 \tag{2.130}$$

with the bond lengths r and the bending angles θ. Here r_0 and θ_0 denote the mean values.

2.1.7 Macromolecules in solution

Important Concepts: Flory-Huggins theory, theta-temperature, phase transitions

The *Flory-Huggins theory* of polymer solutions is based on a lattice theory. It is a mean-field theory.