Modeling Chromatin

Alexis Chan and Jon Harter

University of North Carolina, Chapel Hill

October 18, 2010

Overview

- Defining Chromatin
- Models for Chromatin Ordered by Increasing Complexity
  - Discrete Chain
  - Worm-like Chain
  - Beads on a Wire
- Forces applied to models
- Solvers for Simulation
  - Metropolis-Monte-Carlo
  - Brownian Dynamics
  - Langevin Equation

Chromatin - Definition [1]

- Resides in eukaryotic cells (cells with complex structures contained in membranes)
- Consists of DNA, histone, and other proteins that comprise chromosomes.
- In chromatin, DNA wraps around histone proteins to form nucleosomes.

Chromatin - Function and Structure [4]

- Packages DNA into a small, storable volume within a cell
- The inside of the chromatin fiber and packing of the nucleosomes remains a mystery
- Many hypothetical models attempt to mimic the behavior of chromatin in order to gain insight into its structure

Chromatin - Molecular Structure

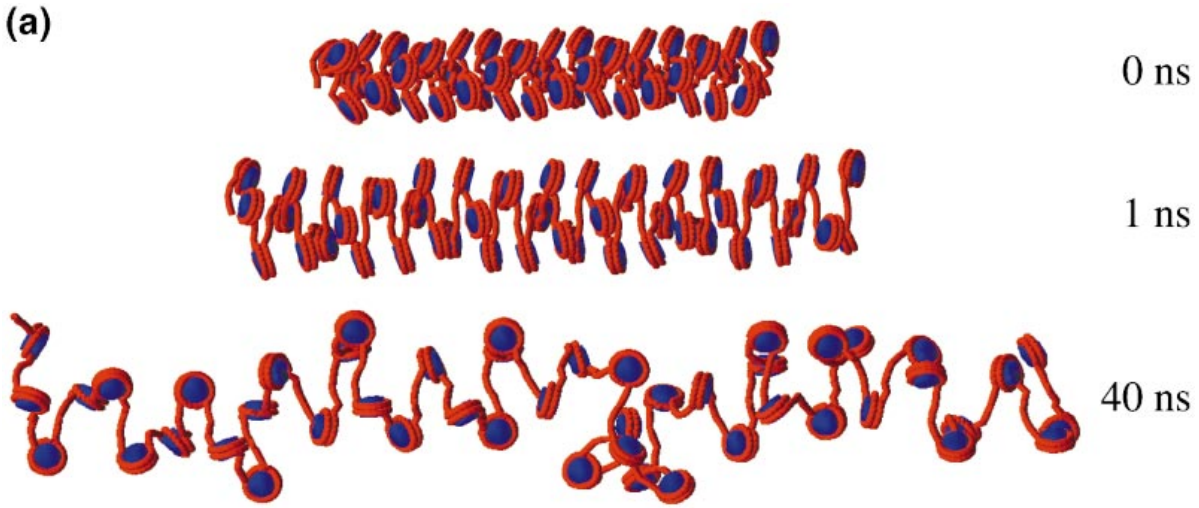


Figure: Chromatin can change in diameter due to stretching and coiling. Above are several different orientations (left) of the same chromatin strand at different times (right) [4]

Chromatin - Molecular Structure

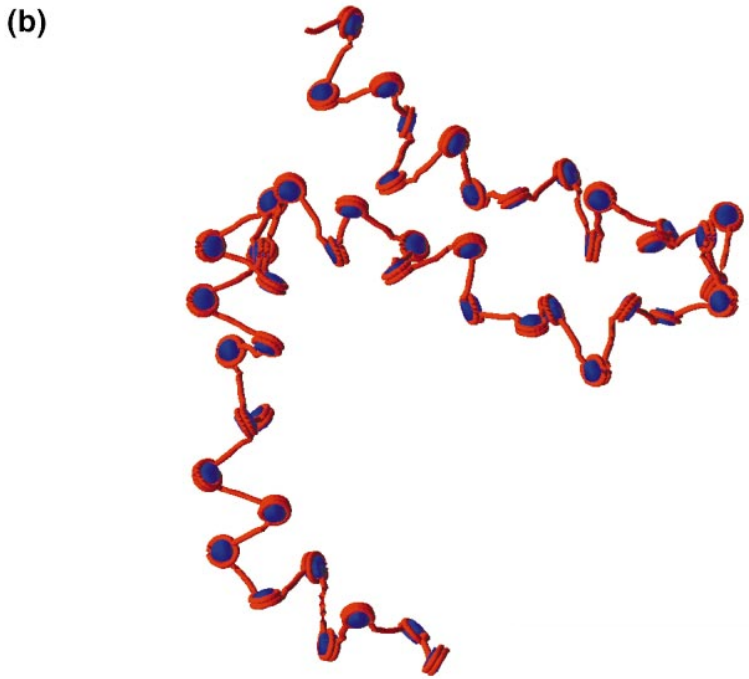


Figure: Example of Chromatin fully unwound [4]

Why Model?

- DNA modeling and simulation is important because it helps us understand the cell behavior and the biological and chemical interactions within cells.
- Because DNA consists of different forms in different phases of a cells life and its actual form can not be measured accurately using current technology, we need to use models to approximate its behavior.

Why Model?

- The complexity of the models we use depends on the biological questions we are trying to solve.
- There is a trade-off between the level of detail we can include in the simulation and the scale of the simulation.
- At higher levels, polymer models of DNA approximate the geometrical structure of the entire DNA within the nucleus and the interaction between DNA strands.
- At lower levels, molecular models of DNA more accurately describe the interactions between the proteins and DNA strands.

Overview - Models

- Discrete Chain - Chromatin under low forces
- Worm-like Chain - Chromatin under high forces
- Beads on a Wire - Incorporates forces from Nucleosomes

Chromatin - Polymer Models

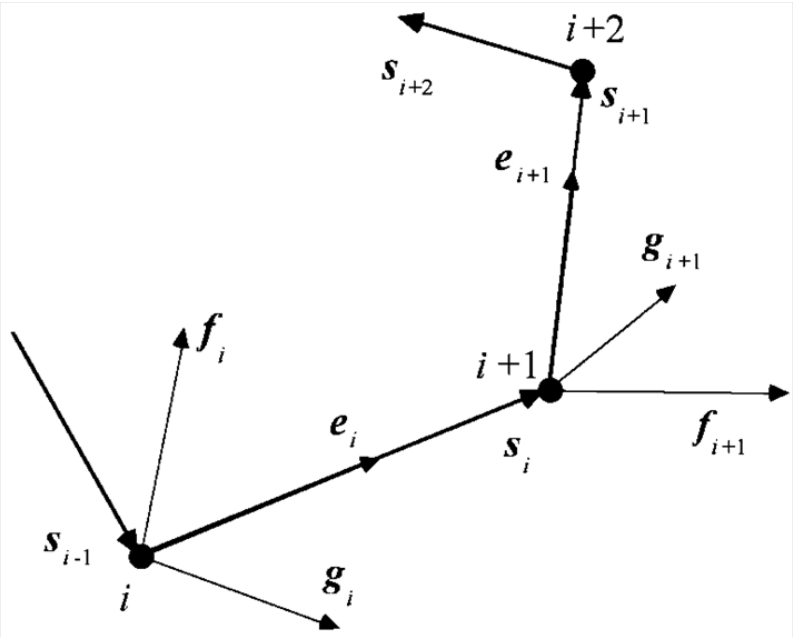
- DNA may be approximated by a flexible elastic chain. An approximate description for such a molecule is constructed by defining suitable segments, which behave like rigid cylinders on the time and length scale considered.

Two types of elastic chain:

- Worm-like Chain
- Freely-jointed Chain

Chromatin - Polymer Models

- The vector s_i defines the direction and length of segment i, f_i is a unit vector normal to the segment and g_i is an auxiliary vector that is used to take into account permanent bending of the DNA. [7]



Chromatin - Polymer Models

- A segmented chain can be described with the following parameters:
  - Bending rigidity
  - Torsional rigidity
  - Stretching rigidity
  - The interaction potential between chain segments

Freely-jointed Chain

- DNA under low external forces can be modeled using the freely-jointed chain.
- Freely-jointed chain: no correlation between the orientation of a segment with the orientation of its neighboring segments
- Chain adopts a random-walk behavior. [5]
- In the low force limit the model acts as a Hookean Spring
- In the high force limit the model reaches full extensibility. This is not the behavior of DNA under high external forces as individual links are extensible. [10]

Freely-jointed Chain

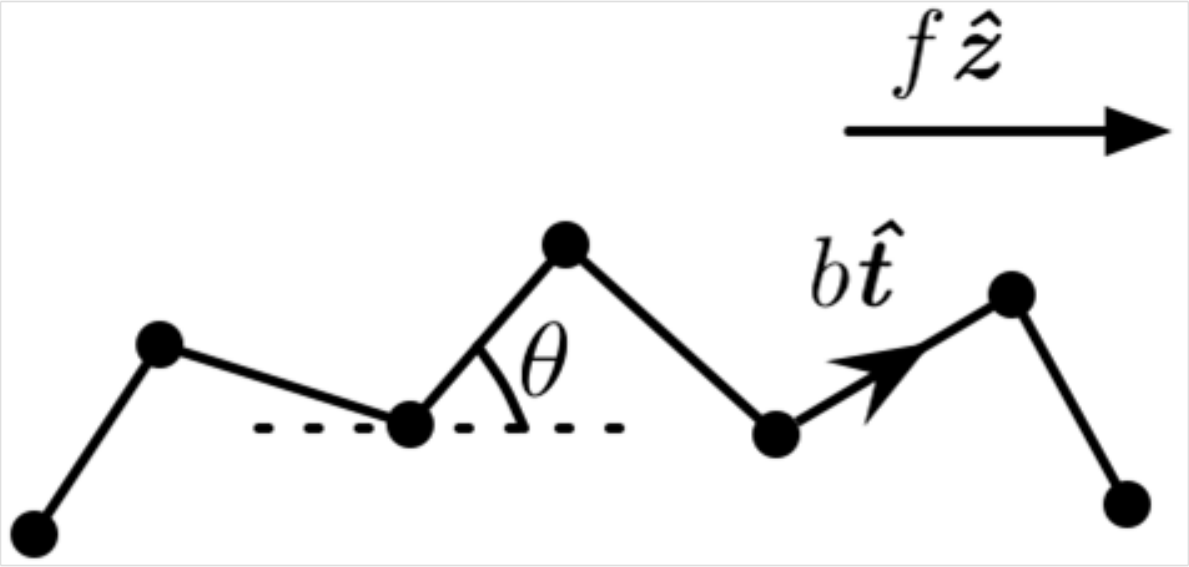


Figure: http://biocurious.com/2006/07/04/wormlike-chains

Worm-like Chain

- The worm-like chain (WLC) equation is commonly used to describe the mode of motion of polymers like DNA [8]
- Basic concept: Amount of force per length exerted by a polymer is proportional to the extension of the polymer

Worm-like Chain

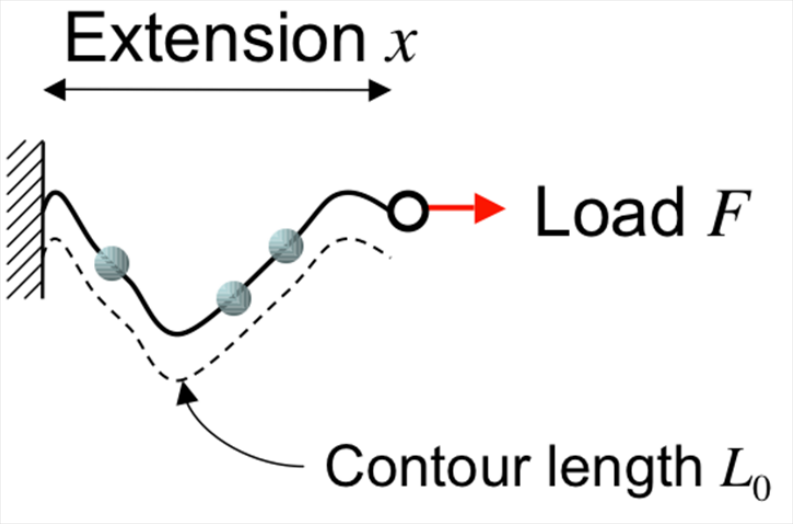


Figure: http://cms.mse.berkeley.edu/research/hierarchical.htm

Persistence Length

Persistence length of a material

- When dividing a chain into segments, the length of the segments matter.
- Segments shorter than persistence length behave approximately like a rigid rod while segments longer than persistence length shows significant internal flexibility. [7]
- The distance over which the correlation of the direction of the two ends of a polymer is lost. [5]
- Describes a filaments resistance to thermal force

Worm-like Chain

Force Exerted by Polymer

$$F = \left(k_B \frac{T}{L_p}\right)\left[\frac{1}{4}\left(1 - \frac{x}{L_c}\right)^{-2} - \frac{1}{4} + \frac{x}{L_c}\right]$$

- F: Force per length exerted by polymer
- x: Extension of polymer
- L_p: Persistence length
- L_c: Contour length - total length of polymer
- k_B: Boltzmann's constant
- T: Temperature

Beads on a Chain - Structure [6]

Definition

- Discrete model used in polymer dynamics
- Can model linear strands of chromatin like wormlike chain, but now allows for the inclusion of nucleosomes
- Each nucleosome is represented as a bead on the Chromatin chain

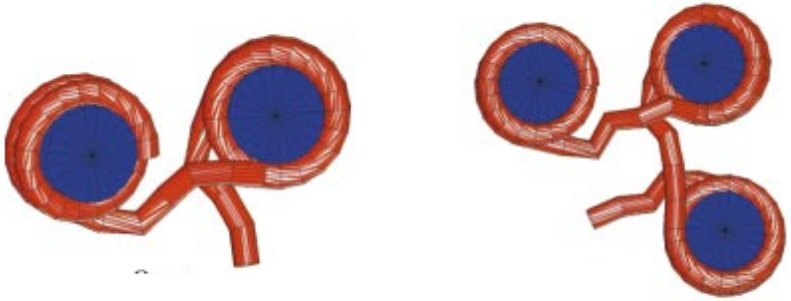
Structure

- N beads of radius a and mass m linked by $N-1$ virtual bonds
- Each bead is linked by an equilibrium "bond length", l_0
- l_0 varies according to the salt concentration of the solution

Beads on a Chain - Higher Detail [6]

Details

- Can add forces directly to beads to describe the movement of the chain
- Can also apply forces directly to the chain
- Simulations may inspect local structures such as dinucleosomes (left) and trinucleosomes (right) [4]



Beads on a Chain - Model Diagram

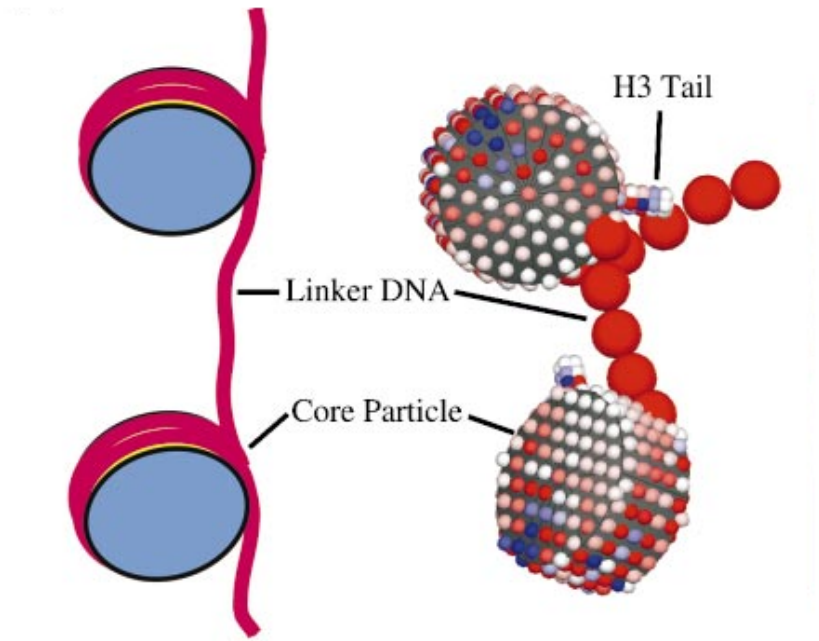


Figure: Depicts nucleosomes (beads) and distributed electrostatic charges (dots on beads) [2]

Overview - Particle Forces

Energy Potentials

- Stretching
- Bending
- Torsional

Forces

- Electrostatic
- Microinic
- Internucleosomal

Young's Modulus [9]

- Stress: ratio of force acting normally to a surface over the area of the surface
  $\sigma_{avg} = \frac{F_n}{A} \approx \sigma$
- Strain: Dimensionless ratio of change in length over original length
  $\epsilon = \frac{L-l}{l} = \frac{\Delta l}{l}$
- Young's Modulus: Normal stress divided by linear strain.
  - Stress/Strain

Persistence Length

Persistence length of a material

- When dividing a chain into segments, the length of the segments matter.
- Segments shorter than persistence length behave approximately like a rigid rod while segments longer than persistence length shows significant internal flexibility. [7]
- The distance over which the correlation of the direction of the two ends of a polymer is lost. [5]
- Describes a filaments resistance to thermal force

Persistence Length

Equation

$L_p = \frac{EI}{k_B T}$

E: Young's modulus

I: Moment of inertia of plane area

k_B: Boltzmann constant

T: room temperature (Kelvin)

$k_B T$: the relevant energy scale for all molecular interactions inside a cell

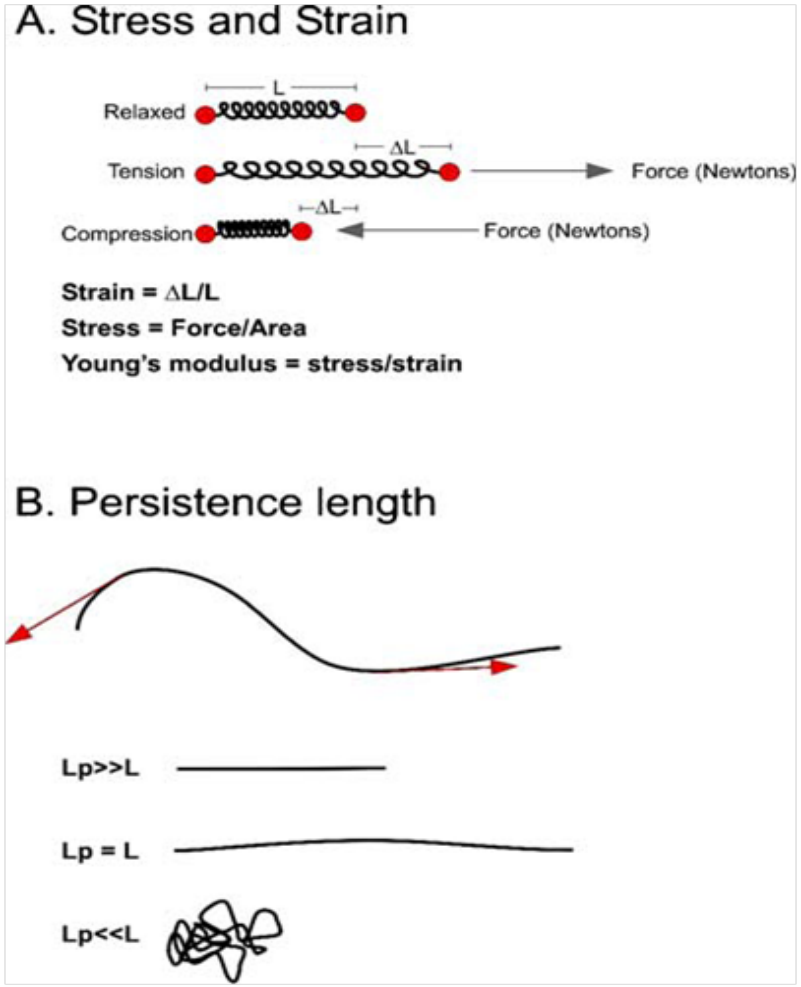
Persistence Length

- The correlation length of the direction of the chain measured along its contour [7]

$$\left\langle \vec{u}(s)\vec{u}(s+s') = e^{-s/L_p} \right\rangle$$

- Here $u(s)$ is a unit vector in the direction of the chain. s is the position along the chain contour, the angular brackets indicating the average over all positions and chain conformations.

Figures - Stress and Strain



Bending Rigidity

Bending elasticity

- The energy required to **bend** a polymer segment of unit length over an angle of 1 radian [7]

- Energy required to bend two segments of length l by an angle θ with respect to one another:

$$E_b = \frac{k_B T}{2} \frac{L_p}{l} \theta^2$$

Torsional Rigidity

Torsional Rigidity

- the energy required to twist a polymer segment of unit length over an angle of 1 radian [7]

- Energy required to twist two segments of length l by an angle ϕ with respect to one another:

$$E_T = \frac{k_B T}{2} \frac{L_T}{l} \phi^2$$

Stretching Rigidity

Stretching Elasticity

- The energy required to **stretch** a polymer segment of unit length

- Energy required to stretch a segment of the chain of length l:

$$E_{str} = \frac{1}{2}\frac{\sigma}{l}\Delta l^2$$

where $\sigma = F * \frac{L}{\Delta L}$ [7]

Intra-chain Interactions

- Intra-chain interactions can be described using the electrostatic potential.
- At low ionic concentrations: Debye-Huckel equation can be used to calculate the elecotrostatic potential between two uniformly charged non-adjacent segments [7]
- Debye-Huckel is a linearized version of Poisson-Boltzmann (explained later).

Intra-chain Interactions

Equation

$$E_{ij}^{(e)} = \frac{v^2}{D} \int \int d\lambda_i d\lambda_j \frac{e^{-\kappa r_{ij}}}{r_{ij}}$$

- λ_i, λ_j: distances from the segment beginnings
- r_{ij}: distance between the current positions at the segments to which the integration parameters
- κ: the inverse of the Debye length ($\kappa^2 = \frac{8\pi e^2}{k_b TD}$)
- I: ionic strength
- e: proton charge
- D: dielectric constant of water
- v: linear charge density

Overview - Solvers

- Metropolis-Monte-Carlo
- Brownian Dynamics
- Langevin Equation

Metropolis-Monte-Carlo Simulation

Basic Idea

- Probability of a particular configuration within a set of N configurations can be estimated using a probability distribution function as N approaches infinity

$$\lim_{N \to \infty} \frac{N_\epsilon}{N} = P(\xi)$$

Metropolis-Monte-Carlo Simulation

Algorithm [3]

1. Pick a configuration
2. Pick a trial configuration
3. Compute the probability ratio R: $P(trialconfiguration)/P(configuration)$
4. Pick a random number p with value between 0 and 1.
5. If $p \leq R$, make the next configuration = trial configuration
6. If $p > R$, next configuration = current configuration
7. Repeat step 2-6 for N times. N has to be a sufficiently large number.

Metropolis-Monte-Carlo Simulation

Using Monte-Carlo to simulate configuration of DNA [12]

- Local configuration:
  - Segment A is randomly selected, and the random displacement is applied.
  - Random vector: Probability distribution of direction is uniform across 3 dimensional space. Length is in the range of 0 and upper bound determined by acceptance ratio.

Metropolis-Monte-Carlo Simulation

- Global configuration:
  - After a random segment is selected, the rotation of the arc including the segment and its neighbors is picked randomly.
  - Probability distribution of rotation axis direction is uniform across 3 dimensional space.
  - Angle of rotation is in the range of 0 and upper bound determined by acceptance ratio.
- Acceptance ratio:
  - If total potential energy of trial configuration is lower than the previous configuration, the trial configuration is accepted.

Metropolis-Monte-Carlo Simulation [6]

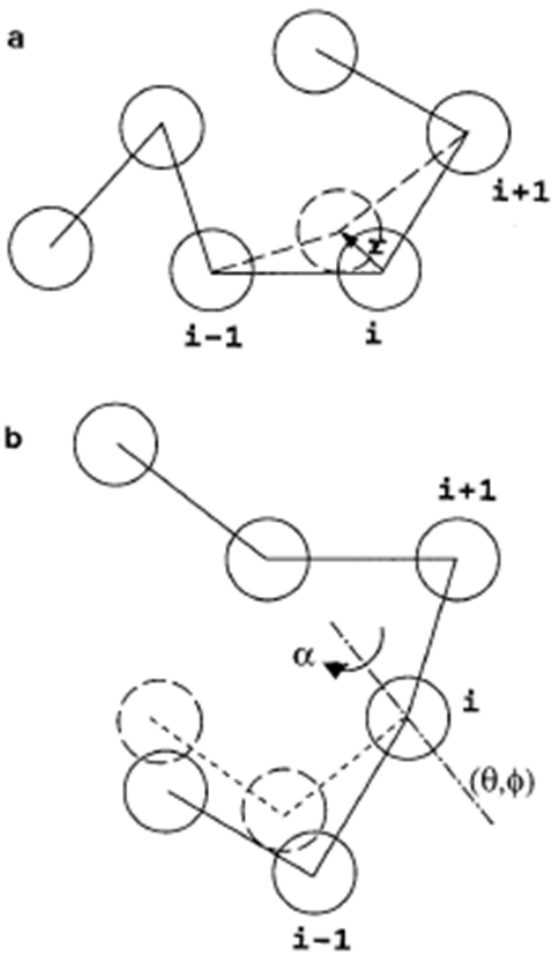


FIG. 1. Two types of Monte Carlo moves: (a) local, random bead displacement and (b) global, random axis rotation.

Brownian Dynamics Algorithm

- First order algorithm by Ermack and McCammon [6]. Generates new position vector from current position vector:

$$\mathbf{x}_{n+1} = \mathbf{x}_n + \frac{\Delta t}{k_B T}\mathbf{D}_n \cdot \mathbf{F}_n + \mathbf{R}_n$$

- $\mathbf{x}_{n+1}$: new position vector
- $\mathbf{x_n}$: old position vector
- Δt: time step
- k_B: Boltzmann constant
- T: Temperature
- $\mathbf{D}_n$: Translational diffusion matrix
- $\mathbf{F}_n$: Collective vector of interparticle forces
- $\mathbf{R}_n$: Random force generated from Gaussian distribution of mean and zero variance ($\langle \mathbf{R}_n \mathbf{R}_m \rangle = 2\mathbf{D}_n \Delta t \delta_{nn}$)
- δ_{nn}: Kronecker delta symbol

Brownian Dynamics Algorithm

- Second order algorithm by Iniesta and Garcia de la Torre. [6]
- 1st order BD and 2nd order BD is analogous to the Runge-Kutta 2 and Euler Method in numerical differential equations.

Brownian Dynamics Algorithm

- Steps:
  1. Perform first order Brownian Dynamics algorithm to determine intermediate position vector x_{n+1}
  2. Compute new position vector using:

$$\mathbf{x}_{n+1} = \mathbf{n}_n + \frac{1}{2}\frac{\Delta t}{k_B T}\left(\mathbf{D}_n \cdot \mathbf{F}_n + \mathbf{D}_{n+1} \cdot \mathbf{F}_{n+1}\right) + \mathbf{R}_{n+1}$$

$$\langle \mathbf{R}_{n+1}\mathbf{R}_{n+1} \rangle = (\mathbf{D}_n + \mathbf{D}_{n+1})\Delta t, \langle \mathbf{R}_{n+1} \rangle = 0$$

  - $\mathbf{D}_{n+1}$: Translational diffusion matrix calculated using $\mathbf{x}_{n+1}$
  - $\mathbf{F}_{n+1}$: Collective vector of interparticle forces calculated using $\mathbf{x}_{n+1}$
  - $\mathbf{R}_{n+1}$: Random force calculated using $\mathbf{x}_{n+1}$

Brownian Dynamics Algorithm

- Brownian dynamics algorithm is a simplification of the Langevin equation. In general, larger time steps are used for Brownian dynamics than Langevin equation. [11]
- At time-scale much larger than momentum relaxation time, we can ignore inertia.
- Specifically:

  $$\Delta t >> m\frac{D_{ii}}{k_B T} = \frac{m}{6\pi\eta a}$$

  - D_{ii}: i^{th} diagonal entry of D
  - m: mass of segment

Translational Motion Solver [6]

Langevin Equation

$$m\dot{\mathbf{v}}_i = -\sum_{j=1}^{N} \xi_{ij} \cdot \mathbf{v}_i + \mathbf{F}_i + \sum_{j=1}^{N} \alpha_{ij} \cdot \mathbf{f}_j, 1 \leq i \leq N$$

- $\mathbf{v}_i$: velocity of the ith bead
- $\mathbf{F}_i$: sum of interparticle forces acting on segment or bead i (model dependent)
- $\sum_j \alpha_{ij} \cdot \mathbf{f}_j$: randomly fluctuating force
- ξ_{ij}: translational friction tensor

Conclusion

- We presented three models of DNA. On the polymer level, we presented the freely-jointed chain and the worm-like chain. On the molecular level, we presented the beads-on-a-chain model. The different levels of modeling are not mutually exclusive you can add levels of complexity to the your basic model. E.g. beads-on-a-chain model can be added to the worm-like chain model.
- We also presented three different methods of simulating DNA movement, in increasing order of complexity: the Metropolis Monte-Carlo method, the Brownian Dynamics Algorithm method and the Langevin Equation method. Once again, the level of complexity that we choose for the simulation depends on the questions we are trying to solve with the simulation.

- Isaac Asimov, *The genetic code*, The New American Library, 1962.
- Frank Aumann, Filip Lankas, Maïwen Caudron, and Jörg Langowski, *Monte carlo simulation of chromatin stretching*, Phys. Rev. E **73** (2006), no. 4, 041927.
- Victor S. Batista, *Metropolis monte carlo method*, `http://xbeams.chem.yale.edu/~batista/vaa/node42.html`.
- Daniel A. Beard and Tamar Schlick, *Computational modeling predicts the structure and dynamics of chromatin fiber*, Structure **9** (2001), 105–114.
- K.S. Bloom, *Beyond the code: the mechanical properties of dna as they relate to mitosis*, Chromosoma **117** (2008), 103–110.

- Hongmei Jian and Alexander V. Vologodskii, *A combined wormlike-chain and bead model for dynamic simulations of long linear dna*, Journal of Computational Physics (1997), no. 136, 168–179.
- J. Langowski, *Polymer chain models of dna and chromatin*, Eur Phys J E Soft Matter **19** (2006), 241–249.
- ME Larson, BD Harrison, and K Bloom, *Uncovering chromatin's contribution to the mitotic spindle: Applications of computational and polymer models*, Biochimie (2010).
- A. D. McNaught and A. Wilkinson, *Compendium of chemical terminology, 2nd ed. (the "gold book")*, Blackwell Scientific Publications, Oxford, 1997.
- PhilipJ, *Worm-like chains*, `http://biocurious.com/2006/07/04/wormlike-chains`.

T. Schlick, *"molecular dynamics: Further topics" ch. 13 in molecular modeling and simulation*, New York: Springer-Verlag, 2002.

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