Monte Carlo simulations of densely packed biopolymers

Outline:

Motivation: dense biopolymers are ubiquitous. Prototypical example: DNA packaged inside a viral particle.

Implications of self- and mutual polymer entanglement?

Methodological aspects: Monte Carlo techniques
Generalities on DNA packaging

**Eukaryotes:** *meters* of DNA in a *10 micron* size nucleus

**Bacteria:** *mm* of DNA in a *micron* size cell

**Phages** *microns* of DNA in a *50 nm* capsid

In all cases genome organisation involves a high degree of spatial confinement.

Jiang et al., *Nature* 2006

Smith et al., *Nature* 2001

DNA packaging time: 5.5mins
What about the DNA arrangement?

**Spooling model,**
Richards et al, JMB 1973

**Fold model,**
Richards et al. JMB 1973

... disordered packing etc.
Imaging studies of DNA in bacteriophages

Phages \( \text{microns} \) of DNA in a 50 \( \text{nm} \) capsid

Cryo-EM imaging on bacteriophage \( \varepsilon 15 \) indicate that the outer layers of dsDNA have an inverse spool arrangement.


Richards et al. JMB 1973
Self-avoidance and bending rigidity

Growth of a flexible self-avoiding chain in a small sphere \((L_p \gg R)\)

See also:
Harvey and coworkers: Biopolymers 73 (2004); Biophys. Chem (2002)
Packing a stiff chain

LaMarque et al., Biopolymers (2004); Arsuaga et al., Biophys. Chem. (2002).

See also:
Harvey and coworkers: Biopolymers 73 (2004); Biophys. Chem (2002)
Knots as a probe of DNA organization

Arsuaga et al. PNAS (2005)

P4 bacteriophage
DNA length: 10 kb ~ 3.4 µm
Capsid diameter: ~ 45 nm

Unknots <5%
Knots as a probe of DNA organization

P4 bacteriophage
DNA length: 10 kb ~ 3.4 µm
Capsid diameter: ~ 45 nm
Ref: Arsuaga et al. PNAS 2005

Some of these knots occur in proteins too!
see Wallin et al. J. Mol. Biol. 2007
Model for circular DNA

Flexible ring of $N$ cylinders
(Vologodskii and coworkers)

Diameter of cylinders: 2.5 nm
Contour length: $N \alpha = 3.4 \mu m$
Persistence length: 50 nm

$$\mathcal{H} = \sum_i V_{br}(i) + \sum_{i \neq j, j \pm 1} V_{hc}(i, j)$$

$$V_{br}(i) = \begin{cases} 0 & \text{if } d_{ij} > 2.5 \text{nm} \\ \alpha(1 - \vec{t}_i \cdot \vec{t}_{i+1}) & \text{if } d_{ij} \leq 2.5 \text{nm} \end{cases}$$

Monte Carlo sampling

MC is used to produce a sequence of system snapshots sampled with canonical weight. Key prescriptions:

1. at each time step obtain a trial system configuration by changing the current one using random moves.
2. Accept the trial configuration or retain the current one using a suitable rule. The accepted/retained configuration becomes the new system configuration.

(1) Monte Carlo moves for polymer chains

end move

crankshaft move
Monte Carlo sampling

MC is used to produce a sequence of system snapshots sampled with **canonical weight**. Key prescriptions:

1. at each time step obtain a trial system configuration by changing the current one using random moves.
2. Accept the trial configuration or retain the current one using a suitable rule. The accepted/retained configuration becomes the new system configuration.

(1) Monte Carlo moves for polymer chains

- **slithering move**
  - Original configuration
  - New configuration

- **pivot move**
  - Original configuration
  - New configuration
Monte Carlo sampling

MC is used to produce a sequence of system snapshots sampled with canonical weight. Key prescriptions:

(1) at each time step obtain a trial system configuration by changing the current one using random moves.

(2) Accept the trial configuration or retain the current one using a suitable rule. The accepted/retained configuration becomes the new system configuration.

(2) Acceptance-rejection rule

We wish that in the long run, configurations are picked with canonical probability

\[ P_{eq}(\Gamma) \propto e^{-E(\Gamma)/K_B T} \]

This condition is satisfied if the rates of going from configuration A to B (and vice versa) obey the detailed balance prescription:

\[ P_{eq}(\Gamma_A)W_{A\rightarrow B} = P_{eq}(\Gamma_B)W_{B\rightarrow A} \]
Monte Carlo sampling

MC is used to produce a sequence of system snapshots sampled with canonical weight. Key prescriptions:

1. At each time step, obtain a trial system configuration by changing the current one using random moves.

2. Accept the trial configuration or retain the current one using a suitable rule. The accepted/retained configuration becomes the new system configuration.

(2) Acceptance-rejection rule

We wish that in the long run, configurations are picked with canonical probability

\[ P_{eq}(\Gamma) \propto e^{-E(\Gamma)/K_B T} \]

This condition is satisfied if the rates of going from configuration A to B (and vice versa) obey the detailed balance prescription:

\[
\frac{W_{A \rightarrow B}}{W_{B \rightarrow A}} = \frac{P_{eq}(\Gamma_B)}{P_{eq}(\Gamma_A)} = \frac{e^{-E_A/K_B T}}{e^{-E_B/K_B T}} = e^{-(E_A-E_B)/K_B T}
\]
Monte Carlo sampling

MC is used to produce a sequence of system snapshots sampled with canonical weight. Key prescriptions:
(1) at each time step obtain a trial system configuration by changing the current one using random moves.
(2) Accept the trial configuration or retain the current one using a suitable rule. The accepted/retained configuration becomes the new system configuration.

(2) Acceptance-rejection rule

We wish that in the long run, configurations are picked with canonical probability

$$P_{eq}(\Gamma) \propto e^{-E(\Gamma)/K_B T}$$

This condition is satisfied if the rates of going from configuration A to B (and vice versa) obey the detailed balance prescription:

$$\frac{W_{A \rightarrow B}}{W_{B \rightarrow A}} = \frac{e^{-E_A/K_B T}}{e^{-E_B/K_B T}}$$

$$W_{A \rightarrow B} = \begin{cases} 1 & \text{if } E_B < E_A \\ e^{-(E_B - E_A)/K_B T} & \text{otherwise} \end{cases}$$
Monte Carlo sampling

MC is used to produce a sequence of system snapshots sampled with **canonical weight**. Key prescriptions:

1. at each time step obtain a trial system configuration by changing the current one using random moves.
2. Accept the trial configuration or retain the current one using a suitable rule. The accepted/retained configuration becomes the new system configuration.

**(2) Acceptance-rejection rule**

We wish that in the long run, configurations are picked with canonical probability

\[ P_{eq}(\Gamma) \propto e^{-E(\Gamma)/K_B T} \]

This condition is satisfied if the rates of going from configuration A to B (and vice versa) obey the detailed balance prescription:

\[ \frac{W_{A\to B}}{W_{B\to A}} = \frac{P_{eq}(\Gamma_B)}{P_{eq}(\Gamma_A)} = \min(1, e^{-(E_B - E_A)/K_B T}) \]
Monte Carlo sampling

MC is used to produce a sequence of system snapshots sampled with **canonical weight**. Key prescriptions:

1. at each time step obtain a trial system configuration by changing the current one using random moves.
2. Accept the trial configuration or retain the current one using a suitable rule. The accepted/retained configuration becomes the new system configuration.

**Cons:**
- No viable information about system kinetics. However, if one uses only **local** moves, then MC trajectories can be a viable stochastic system dynamics.

**Pros:**
- Efficient exploration of phase space
- Information about system equilibrium properties
- Potential energy needs not be differentiable
- Constraints can be efficiently implemented
Self-avoidance and bending rigidity

Growth of a flexible self-avoiding chain in a small sphere \((L_p \gg R)\)

See also:
Harvey and coworkers: Biopolymers 73 (2004); Biophys. Chem (2002)
Advanced sampling techniques: parallel tempering

Sampling the relevant phase space is impractical due to large (free) energy barriers.

How can we overcome the problem?
Advanced sampling techniques: parallel tempering

Run several MC trajectories at additional (higher and lower) temperatures and occasionally propose swaps between systems at nearby temperatures.
Advanced sampling techniques: parallel tempering

Run several MC trajectories at additional (higher and lower) temperatures and occasionally propose swaps between systems at nearby temperatures.

Q: With what probability should we accept the swap?
Advanced sampling techniques: parallel tempering

Run several MC trajectories at additional (higher and lower) temperatures and occasionally propose swaps between systems at nearby temperatures.

Q: With what probability should we accept the swap?

\[
\min \left(1, \frac{e^{-E_B/T_A} - E_A/T_B}{e^{-E_A/T_A} - E_B/T_B}\right)
\]
Stochastic sampling of compact rings

Use Metropolis scheme to sample rings with weight

\[ w \equiv e^{-P R^3 - \beta H} \]

Rings are deformed by crankshaft moves


Recover canonical statistics by undoing pressure bias (Ferrenberg and Swendsen, PRL 1989)

Maximum packing: R ~ 50nm
Confinement and knot complexity
### Simple knots

#### Unconstrained case:

<table>
<thead>
<tr>
<th>Knot type</th>
<th>Probability</th>
<th>Experiment (*)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>3.8%</td>
<td>3.5%</td>
</tr>
<tr>
<td>4</td>
<td>0.46%</td>
<td>0.44%</td>
</tr>
<tr>
<td>5</td>
<td>0.27%</td>
<td>0.25%</td>
</tr>
</tbody>
</table>

* 1M NaCl

Rybenkov et al. PNAS 1993
Packing of half P4 genome (4.7Kb)
[experiment: Trigueros and Roca BMC biotech. 2007]

1. No order at surface
2. No bias in favour of torus and chiral knots

Configuration obtained with “growth” simulations (kink-jump dynamics)

What is the missing ingredient?
Dense phases of DNA segments

adapted from R. Podgornik, Taiwan lectures

Estimated P4 DNA density (packaging model of Purohit et al. PNAS 2003)

Full genome (10kb) : 270 mg/ml
Half-genome (4.7kb): 200 mg/ml
**Cholesteric phases of DNA**

DNA strands form a preferential angle (steric hindrance + electrostatics)

Leforestier et al. C. Rendu Chimie (2008)
Ferrarini et al. J Chem. Phys (2005);

Introduce additional cholesteric potential (besides chain connectivity, bending energy and screened electrostatic interactions):

\[
V = k(\alpha - \alpha_0)^2 f(d_{i,j})
\]

\[
\tan(\alpha) = \frac{(\vec{b}_1 \wedge \vec{b}_2) \cdot \vec{d}_{12}}{(\vec{b}_1 \cdot \vec{b}_2) |\vec{d}_{12}|}
\]

\[
n = \frac{1}{2.5 \text{nm}}
\]

\[
\theta = 1^\circ
\]

\[
k \approx K_B T
\]
Ordering effect of the cholesteric potential
Ordering effect

Marenduzzo et al., PNAS, 2009
Knot spectrum (after circularization)

\[ \alpha_0 = 1^\circ \]
\[ k = 2 \ K_B \ T \]

\[ \alpha_0 = 10^\circ \]
\[ k = 1.0 \ K_B \ T \]

Bias towards torus and chiral knots over a good range of parameters
Potential strength tuned to reproduce experimental data on full P4 genome

Knots are delocalised; on average they occupy 60% of the chain.
Ejection of entangled DNA
... and one more thing: thermodynamic reweighting

MC simulations at a given temperature can give us equilibrium properties at different (nearby) temperatures!

The height of the $i$th bin is proportional to:

MC time series

<table>
<thead>
<tr>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>time</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>histogram</td>
</tr>
</tbody>
</table>

E
thermodynamic reweighting

MC simulations at a given temperature can give us equilibrium properties at different (nearby) temperatures!

The height of the $i$th bin is proportional to:
- $N$, the total number of MC snapshots
- $\exp(-E_i/T)$, the canonical weight
- $W_i$, The number of microstates with energy $E_i$

$$h_i \propto N W_i e^{-E_i/T}$$
thermodynamic reweighting

MC simulations at a given temperature can give us equilibrium properties at different (nearby) temperatures!
... and one more thing: thermodynamic reweighting

MC simulations at a given temperature can give us equilibrium properties at different (nearby) temperatures!

\[
\langle E(T') \rangle = \frac{\sum_i E_i W_i e^{-E_i/T'}}{\sum_i W_i e^{-E_i/T'}}
\]

\[
W_i \propto \frac{h_i}{N} e^{+E_i/T}
\]
Summary

- Monte Carlo as a general tool to characterize equilibrium properties of systems
- Advanced sampling techniques: Parallel tempering
- Thermodynamic Reweighting techniques

Application to a challenging system: densely packed DNA

Useful references (based on my own taste…):

- Itzykson & Drouffe, Statistical field theory
- Newman and Barkema, Monte Carlo methods in Statistical Physics
- K. Binder, Lecture notes of Varenna summer school
- + material available at http://people.sissa.it/~michelet/Lund
Acknowledgements

- Pietro Faccioli (UniTN)
- Davide Marenduzzo (Edinburgh)
- Ken Millett (UCSB)
- Enzo Orlandini (Padova)
- Raffaello Potestio (Mainz)
- Angelo Rosa (SISSA-IIT)
- Andrzej Stasiak (U. Lausanne)
- Tatjana Skrbic (UniTN)
- De Witt Sumners (FSU)
- Luca Tubiana (SISSA)

- Javier Arsuaga
- Giovanni Dietler
- Eric Rawdon
- Ngo Mihn Toan
Workshop on Physical Virology
(24 - 28 September 2012)
Mineo – Piana, Italy

The Abdus Salam International Centre for Theoretical Physics (ICTP, Trieste, Italy) is organizing a Workshop on Physical Virology to be held at the ICTP from 24 - 28 September 2012.

DESCRIPTION:
Our competitors to bacterial and phylogenetic classification are infectious organisms that achieve a high efficiency of their replication cycle by utilizing various processes and active physical mechanisms.

The Workshop aims to gather researchers from different disciplines to focus on recent theoretical and experimental advancements on the efficiency of these physical mechanisms, particularly those with a biological basis.

The main topics will be:
1. self-assembly and generation of viral capsids
2. packaging and ejection of the viral genome from the capsid
3. movement between the viral RNA and capsid protein
4. virus-host membrane interactions
5. generation of antigenic diversity of membrane proteins

PARTICIPATION:
Academic and research staff of science communities that are members of the Systematic Nomenclature Applications of Taxonomy (SANT) are invited. As the Workshop aims to conduct in English, participants should demonstrate adequate working knowledge of this language. Although the main purpose of the Centre is to help research workers from developing countries, through a program of training activities within a framework of international cooperation, a limited number of candidates and post-doctoral students from developing countries are also welcome to attend.

As a policy, travel and subsistence expenses of the participants chosen by ICTP for their travel arrangements and accommodation expenses will be paid. Travel grants will be available for potential speakers who are nominated by ICTP members and who are not more than 35 years old. Travel grants are available only for those who attend the entire Workshop. There is no registration fee for attending this activity.

DIRECTORS:
Tristan LINDEAUX
ICTP and University of Trieste, Italy

Valérie LORIAN
ICTP and University of Trieste, Italy

Cristian MICHELUTTI
ICTP, Italy

Balaji PODGORNIK
University of Illinois, Urbana-Champaign, USA

LOCAL ORGANIZERS:
Matteo MAZELLI
ICTP, Italy

INVITED SPEAKERS (IN ADDITION):

Shuhei OSHIMA (2pm, Trieste)
Ralph O'BRIEN (Washington, USA)
Ralf SUHRBAER (4pm, Trieste)
Alex ANDERSON (from USA)
Mieke HINNE (from USA)
Flavia DE VITÀ (from USA)
Lyam O'BRIEN (from USA)
M. LAURIA (from Italy)

APPLICATION DEADLINE:
June 20

ICTP Workshop on Physical Virology
Application deadline: June 20